AI Against Misinformation: The Role of Machine Learning for Fake News Detection

Nada Alghamdi, Raghad Aljiban, Kayan Almesned, Shoug Almoaibed, Aseel Alkhaldi, Mozoon Alkhalis, Nora Aljomuh

Department of Computer Science, College of Computer Science and Information Technology, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia

2220005359@iau.edu.sa, 2220004952@iau.edu.sa, 2220004499@iau.edu.sa, 2220004372@iau.edu.sa, 2220006043@iau.edu.sa, 2220002873@iau.edu.sa, 2220004452@iau.edu.sa

Abstract

In an era where the spread of information is real-time, the phenomenon of fake news is a monstrous menace to public trust, well-informed choices, and peace. Various studies have been undertaken on this aspect but keeping pace with the changing and dynamic situation of misinformation—particularly in social media—remained to be addressed by detection mechanisms. Various machine and deep learning architecture types such as Naïve Bayes, Logistic Regression, CNN, and BiRNN have already been used with impressive accuracy. Solutions based on them are, however, typically context- or dataset-specific and are not dynamically scalable to general world environments. This limitation prompts work on more stable, scalable solutions.

This research compares the execution of two "Machine Learning" techniques in detecting misinformation: 'Support Vector Machine" (SVM) and "Random Forest" (RF). SVM is employed here due to its ability to execute productively in high-dimensional feature spaces with linear boundary separation, as well as RF due to its ensemble learning strategy and resistance to overfitting to prevent model instability. Both models were trained and evaluated on labeled data with "accuracy" and "F1-score" as measures.

Outcomes demonstrate that both models yield approximately equal performance. The study contributes to ongoing efforts to enhance the reliability of AI-based misinformation detection and highlights the need for further research into multimodal and hybrid approaches.

Keywords: Fake News Detection, Machine Learning, Classification Algorithms, Support Vector Machine, Random Forest, Model Evaluation, Feature Extraction.

1.0 Introduction

The growth of the social networks and web portals has made distribution easier and quicker than ever. This added convenience comes at the expense of exposed risks—from the phenomenon of false news in the form of fake news, which gets published with an intention to mislead or simply ignored and invented. The duplication is injurious to the reputation of publishers, public perception, and political action. The problem must be tackled by designing computer systems to identify, classify, and discriminate such fake news.

The issue receives extra research emphasis through the application of machine learning (ML) according to [1],[2]. The achievement of rising misleading information detection as a task allows machine learning models to detect fake news through methods including Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Naïve Bayes (NB) and deep learning models CNN and RNN. The approaches seek to identify news articles through specified text input with optional metadata parameters. Such detection solutions face ongoing challenges due to the dynamic changes in misinformation dynamics specifically when occurring within social media platforms. Due to their inability to generalize across different datasets and platforms effectively these models require new enhanced detection frameworks according to [4].

Regardless of all the research conducted, the broad spectrum of focusing either on tailored datasets or a lack of integration for forming strategies from fakes content generators is a problem. [3] illustrates the drawback of using too stringent models to practical frameworks; even the highest precise ones, at 99.9% for Stacking and 99.8% for CNN and BiRNN, are plagued by overfitting or reduced performance on different platforms or live environments. In addition, there is no study that includes a model comparison where all the participants are placed in controlled same conditions in the same situation to examine consistency. Therefore, there is still a need for a study that acts on different ML classifiers under the same real-life situation to verify which model is optimal in terms of stability and flexibility.

Our work of research used two major ML classifiers - Support Vector Machine (SVM) and Random Forest (RF) for identifying fake news. There is the text preprocessing phase prior to TF-IDF feature extraction that yields model training using a labelled data set. After we have searched past work experimented that tried to utilize these models, hence we found that they were suitable to use. The Support Vector Machine classifier performs excellent binary classification of complex text data by constructing optimal maximum-margin classification boundaries.
RF applies ensemble learning to build several decision trees that provide resistance to noise and overfitting and offers high accuracy and interpretability.

We utilized both classification methods and experimented them on the same set of data with uniform test conditions. The accuracy and precision and recall and F1-score metrics were used for comparison purposes. The first result indicates Random Forest as the ideal choice for noisy text data and strong generalization while SVM is excellent for text classification precision.

- ✓ Following the introduction the report follows these organizational structures:

  Section 2 Review of related literature.

  Section 3 Project Deliverables of the team

  Section 4 introduces the suggested machine learning methods which combine Support Vector Machine (SVM) with Random Forest (RF).

  Section 5 presents information about empirical studies by showing details about dataset characteristics together with experiment parameters and evaluation metrics with optimization procedures.

  Section 6 Presents the results and discussion.

  Section 7 Conclusion of the report and recommendations.

2.0 Review of Related Work

A. Literature Literatures

1. Overview

The rise of fake news became a considerable challenge due to the proliferation of information in the digital age. This phenomenon pointed out serious threats to public opinion. In the last five years studies have shown that fake news and misinformation can spread rapidly, influencing behaviors on a large scale. Over the decade, many techniques involving the detection of fake news developed side by side with the expansion of social media users.

This literature review aims to scout about using machine learning (ML) to detect fake news. Some studies highlighted the significance of machine learning to detect misinformation, but ever-changing features pose a challenge to categorizing them. On the other hand, deep learning focusing on hierarchical features, is being put into practice both in research and applications, due to the limitations of machine learning.

1. Related Work

The research conducted by S. I. Manzoor et al. (2019), outlined various methods related to detecting misinformation, which include linguistic Modeling, Deceptive, clustering, Predictive Modeling, a Content cue-based approach, and non-text cue-based methods [5]. Furthermore, many studies have implemented supervised machine learning classifiers to find the real from the fake news, the used model in the research of A. Ahmed et Al. (2021) is SVM (Support Vector Machine), is applied for text classification tasks. Logistic Regression is a classifier that is commonly used in binary classification problems and has been operated for fake news detection (Kaur et al., 2020), another classifier that outperformed traditional methods is Random Forest used for utilizing multiple decision trees and enhances accuracy in pointing out misinformation, reducing overfitting and improving generalization. (Ni et al., 2020). Researchers outlined that classifier selection should consider both computational efficiency and accuracy. Moreover, to get the full effectiveness of the classifiers, they need to be trained. Labeled datasets are split into two sets, training and testing to ensure that models could generalize well to unseen data (Wang et al., 2020). Future studies should take into consideration the enhancement of detection accuracy and adaptability by exploring deep learning and unsupervised approaches [6].

Due to the spread of fake news, people's reactions and perceptions to actual news can be altered by fake news, leading to mistrust and misunderstanding. AI system classification techniques have seen success in categorizing problems, especially with the latest modernization in equipment and increasing dataset size. The accuracy of systems that rely on AI and machine learning for deception detection is more or less 70%. Uma Sharma et al. (2021) studied a fairly straightforward approach for detecting fake news through the naïve Bayes classifier. Fake news according to them is usually characterized as relying heavily on emotion and grammatical errors while also seeking to manipulate readers' opinions on an issue.

In the paper the system was explained in three sections. The first section is fixed which runs a machine learning classifier. Model was trained with four classifiers, and the preferable one was selected for ultimate run. The second section is dynamic, which takes the text/keyword given by the end-user, it investigates online for the truth likelihood of the news. The last section verifies the validity of user URL input. The challenge of false news makes the use of digital technologies difficult and can influence people to be convinced of the truth of their opinions. However, machine learning and NLP methods is used to create a fake news labeling system. After that the model is trained on an appropriate dataset, with many metrics being employed to validate its performance. The best model for static search was Logistic Regression with 65% accuracy. Then, using the grid search parameters was optimized and managed to get 75% accuracy, which means that there was a 75% probability that the data would be classified correctly. The accuracy of the dynamic system increases to 93% every iteration. Web Crawler and Online databases will be used to collate all live news and data in a database and create a dataset that will be updated with recent news [7].

As discovered during the research by Jamal Abdul et al., implemented in 2021, much like in the previously discussed contexts, early attempts of developing algorithms to detect fake news relied on traditional machine learning methods. Numerous news classification works were completed using decision trees, random forests, and support vector machines. While basic, these methods had a high reliance on feature engineering, which severely limited their efficiency and flexibility across different datasets (Wang, 2017; Ma et al., 2016). The feature extraction process was done manually, which not only made the entire process tedious but also raised bias depending on the chosen features.

The introduction of deep learning changed the paradigm of fake news detection. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) emerged as potent text classification tools. CNN’s primary strength is in the

recognition of spatial hierarchies in the text, which makes them ideal for feature extraction. RNNs, and more specifically long short-term memory (LSTM) networks, learn temporal dependencies and are very good at dealing with ordered sequences of data (Kollias & Zafeiriou, 2020; Masood et al., 2018). Because of these complementary strengths, deep learning models can surpass the performance of traditional methods in many situations. Novel developments in this area focus on the creation of hybrid models which combine the two most successful approaches, CNNs and RNNs.

Hybrid models with a combination of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been developed recently. Such a solution is an optimization of both architectures and their synchronization that provides a higher quality of feature extraction and a better understanding of the news articles in context. Research presented by Ruchansky et al. (2017) and Yang et al. (2018) has shown that such hybrid techniques can notably improve the detection power in a variety of classification problems. However, there are a few imponderables yet in the domain of fake news detection. One of the issues is called Generalization which means that most models have high performance on specific datasets but fail to generalize to other datasets or even other contexts and data distributions (Vosoughi et al., 2018). This frailty of fixing the models to solve a particular problem can make them useless in case the system is quite big and very different from what is currently known.

Another stumbling block is overfitting, a usual problem in deep learning where the models work very well on the training set of data, but they are unable to predict the results on the unseen test set of data (Drumond et al., 2019). In such circumstances, the question about the trustworthiness of these neural network models as referring to dynamic events characterized by sudden changed information-dependant landscapes should be brought to the forefront. In addition, the accuracy of the detection models can be related to the diversity and size of the

training datasets. The datasets need to be as wide as possible. News data and topics and sources should be represented in numerous genres, and they need to be gathered in a holistic manner for long-term subsistence. For example, work such as Elhadad et al. (2019) argues the point of a diversity of training data collected to show the success of the model by halting the biases. Fourier smaller and less diverse datasets tend to appear [8].

The predominant area of research dealing with clickbait detection has been that of English-language content and the content of this research. For example, Chakraborty et al. (2020) collected 15,000 news articles from Wikinews and similar platforms but indirectly for clickbait issues. The study reached the final inter-annotator agreement, measured by Kappa of 0.79, showing the reliability of the labeling methods. Similarly, Agrawal et al. (2018) tackled the same issue through the development of datasets by utilizing a variety of social networks. The main role of the crowd-sourced annotators in that work was to classify headlines and the success of the collaborative approach in the generation of datasets. Although such milestones exist, Arabic clickbait detection studies have been few and far between. On the one hand, Alzanin and Azmi (2021) handled rumor detection relating to Arabic tweets but on the other hand, they entirely neglected clickbait detection or the creation of exclusive datasets.

This for a huge disparity in research depicts the essence of specially created resources for Arabic content. The study called "Building a Perfect Dataset for Arabic Fake News Detection" authored by Bsoula et al. (2022) is of crucial significance in this respect. The creators offer a groundbreaking establishment for a unique and entirely new dataset ("a dataset that didn't exist before") that is Arabic, a dataset that was created on the basis of the samples drawn from tweets (tweets representing the content of the news) published or tweeted by 24 Jordanian news outlets, out of them 18% are or contains clickbait. The three annotators who

painstakingly labelled each record managed to reach an 81% agreement rate; this was a solid base for subsequent research.

Assessing the usability of a dataset written by the authors was possible by various machine learning algorithms including Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), Naïve Bayes (NB), Stochastic Gradient Descent (SGD), Nearest Neighbor (NN), and Decision Tree (DT). The models' performance was revealed by Using Macro F1-Score, showing the dataset's effectiveness for automatic clickbait detection. The findings are interesting: the first sample data had a Macro F1-Score that ranged from 0.37 to 0.81, and the second sample, which was oversampled, got a range of 0.29 to 0.78, and the third sample, which was undersampled had scores between 0.18 and 0.74. The top F1-Score ratios were registered with the combinations of SGD, LR, and particular attributes, which indicates the ability of the dataset to support powerful machine learning applications in Arabic clickbait detection [9].

As more years pass, traditional detection methods that primarily depend on linguistic analysis are becoming less effective due to the increasing sophistication of AI-generated misinformation. Park and Chai (2023) identified that the current models in detecting misinformation need intelligence much more important than just the analysis of the text. After integrating user behavior, social network dynamics, and content features, a machine-learning technique can work within the process, making it all more adaptable and accurate. The datasets on which the researcher based his work consisted of news articles and social media postings wherein textual features were obtained, as were network features. The researchers used a dataset comprising news articles and social media posts, extracting both textual and network-related features. Key factors influencing fake news detection included: “word sentiment”, “in-degree centrality”, “word similarity”, and “total number of posts”. These features were then fed into several machine learning models, of which the RF (Random Forest) model performed best, attaining an

accuracy of 94.1%. These challenges in future work could be addressed by making the detection systems using machine learning more adaptable and transparent [10].

As Altheneyan and Alhadlaq (2023) discussion were based on engagement of machine learning and big data for pointing fake news via Apace Spark on social media. Through the use of evaluation metrics, the trained model is able to produce a high performance in classification, which yields 92.45% on F1-score. The model achieved the best-performing value of the F1 score and outperformed all of the prior baseline methods. The proposed ensemble model achieves 9.35% improvement in F1 score. Also, one of the limitations to the study model is to Spark, because it requires long training time as it is working independently in a cluster. because of the solitary cluster, Spark takes two times as lengthy to educate. The proposed paintings may be extended for future use with the aid of working many neural community-based fashions, which are more enough for unsupervised fake information detection [11].

Decisional algorithms achieve their quality through accurate measures, but the selected datasets affect these results. Research now exhibits extraordinary accuracy levels reaching above 98.0% in their recent studies. The research evaluates false news identification algorithm precision while introducing better algorithms and datasets. The main objective behind the study will be to generate dialogue about fake news along with emphasizing the significance of computational research toward solving it. The majority of fake news spreads from social media networks to cause political, economic and social effects. Scientists have established detection procedures for fake news with four main categories including clickbait, propaganda, satire and parody, hoaxes, etc. A panel of experts formed by the European Commission actively provides recommendations and discusses proposed measures to prevent fake news distribution along with online disinformation dispersal. The detection of fake content in articles depends on machine learning programs analyzing written information through datasets [12].

According to a study conducted by Villela et al. (2023), seven databases were used, specifically ACM Digital Library and IEEE Xplore. Maximum accuracy was found to be a whopping 99.9% for the Stacking Method, while BiRNN and CNN models reached 99.8% accuracy. Other algorithms such as Gradient Boosting, Long Short-Term Memory, and Robustly Optimized BERT Pretraining Approach for fake news showed an accuracy of 99.5%, 99.4%, and 99.3% respectively. Various datasets are utilized in systems for fake news detection according to RQ2 of this study. Studies have shown the scientific advancement in fake news detection through computational techniques that requires precise recognition methods. The high accuracy scores came from the Stacking Method which achieved 99.9% while BiRNN reached 99.8% with CNN at 99.8% accuracy. The researchers preferred to use CNN algorithms among all other approaches. Among the available datasets Kaggle stood out as the most popular whereas Weibo and FNC-1 and COVID-19 Fake News followed behind it. Future research studies need to examine additional datasets from alternative languages while developing ensemble models and applying deep learning principles to perform fake news detection. The detection capability of computational techniques has improved but scientists require additional research to deal with increasing fake news complexity [12].

Jouhar et al. (2024) has contributed a study "Fake News Detection using Python and Machine Learning" where they tackled this particular issue by considering a range of machine learning algorithms to improve decision-making and to make sure that the acquired knowledge is reliable and accurate. These authors hope to develop a plan that will help them to assess these algorithms systematically in order to select the most effective one. Both algorithmic efficacy measures, such as precision, and accuracy metrics, play the role of factors in model performance which are essential for the use of comprehensive algorithms by the authors to achieve accurate results.

In the previous literature of the fake news detection field, the encountered new things are of several kinds including traditional machine learning methods and

naturallanguage processing (NLP). On the other hand, Sharma et al. (2020) created a set of classifiers (Passive Aggressive Classifier, Random Forest, and Logistic Regression) which were effective in the fake news detection task, but they needed different results in certain cases. Although some cutting-edge models like XGBoost and Decision Trees were the ones employed by Khanam et al. (2020) and Pandey et al. (2021), yet their outcomes were still far from the optimal levels to reach either accuracy or F1-scores. These pieces of the study become arrows that show the board of the question with respect to the issues and values of each method [13].

The literature points out that NLP technologies are useful, yet their major need is in the field of speech, and they still stumble due to subtle contexts and adversarial attacks. At the same time, traditional machine learning techniques are more likely to be biased by the quality and representativeness of the training data. This fact emphasizes the necessity of the adaptive development of modern methods which must be capable of flexibly and swiftly changing along with the rising phenomenon of misinformation transmission. The study by Jouhar et al. calls for a thorough study of the numerous machine learning algorithms, so as to inspire technology to break.

The selection of a qualitative dataset for the training of machine-learning algorithms is a significant part of the study. In the work presented, the authors made use at first of the ISOT fake news dataset, which is very popular for its numerous articles and well-balanced origin of real and false news. The dataset consists mainly of the title, text, and publication date, among other things, and eases the progress of the exploration data analysis (EDA). EDA performs a very important function in the course of revealing data such as trends, correlations, and special features that make a clear distinction between true and fake articles.

The paper also lays significant stress on data preprocessing and feature extraction. The authors used the TF-IDF (Term Frequency-Inverse Document Frequency) method to transform text data into a numeric format, thereby making it

applicable for machine learning. This method involves assessing the importance of terms in a document based on their rarity across the dataset, and thus it provides a background for learning that enables better modeling.

On the lines of the approach, model estimation, Jouhar et al. considered, among other models, Logistic Regression, Decision Trees, Random Forest, Gradient Boosting, XGBoost, and a Passive Aggressive Classifier to evaluate their achievements. The result revealed that the XGBoost algorithm was better than all the other models in terms of accuracy and false positives. This fact was backed up by previous work findings that argued when traditional methods like ensemble ensemble are applied to classification problems the outcomes become superior [13].

B. Gap Identification

Pointing out the gap in the papers related to detecting fake news, after thoroughly reviewing them. Although various machine learning algorithms were tested, like SVM, LR (Logistic Regression), Naïve Bayes, and deep learning approaches such as (CNN and RNN), etc. most of these are confined to either controlled datasets or specific contexts. This in turn limits their applications in real-world scenarios, especially on social media platforms where misinformation runs amok. While several works have foregrounded the importance of using diverse datasets across different languages and cultural contexts, not many works have applied hybrid models or researched how visual information, and multi-modal datasets could be integrated in misinformation detection. Moreover, there is a serious dearth of structured frameworks considering the dynamic nature of fake news and their manifestations across varied platforms-elucidating, in fact, the need for more robust systems that are able to adapt effectively in dynamic sets.

The studies also prove the efficiency of numerous algorithms that cover Support Vector Machines (SVM) and recurrent neural networks similar to LSTM, but the existence of extensive comparisons analyzing the performance of the classifiers on a wide range of data sets and also real-time applications is still a missing point. Besides, the papers show not enough attention is paid to the issues raised by multilingual misinformation, which confines the models proposed to be really universal throughout the world. To an equal extent, there is very little focus on ensemble methods that bring different features of algorithms together on detection accuracy. Research exclusively mentions numerical data from social media, which overshadows the potential of unstructured data sources, e.g. video and audio content, the use of which is becoming a trend in the misinformation landscape. Addressing these gaps would further enhance accuracy and reliability concerning the detection of fake news systems in the dynamically changing digital space.

Ref	Title	Year	Dataset	ML	Result	Notes
5	"Fake News Detection Using Machine Learning approaches: A systematic Review"	2019	**LIAR**: Consists of categorized statements. Fakenewsnet: Consists of news articles labeled as "fake", "biased", or "conspiracy".	- Naïve Bayes - Decision Trees - Support Vector Machine (SVM) - Neural Networks - Random Forest - XGBoost	- An accuracy of various models is noted to be between 63% to 70%.	The authors highlight that existing machine learning approaches have limitations in improving detection accuracy due to the evolving nature of fake news.
6	"Detecting fake news using machine learning: A systematic literature review"	2021	**LIAR**	- SVM - Naïve Bayes - Logistic Regression - RF - K-Nearest Neighbor - Neural Networks - Recurrent Neural Network - Decision Trees - XGBoost	- Various classifiers reported accuracies ranging from 74% to 96.08%, with Naïve Bayes achieving 96.08% in one study. - SVM and LR also noted for their effectiveness in achieving high accuracy.	Other datasets mentioned but not specified in detail.
7	"Fake News Detection Using Machine Learning"	2021	**LIAR**: A benchmark dataset containing 12,836 human-labeled short statements from several contexts. **REAL_OR_FAKE.CSV**: Contains news content with labels indicating whether the news is fake or true.	- Naïve Bayes Classifier - RF - LR - Passive Aggressive Classifier	- The best performing model was **Logistic Regression** after parameter tuning, achieving an accuracy of **80%**. - For the dynamic system using the Passive Aggressive	-

					Classifier, an accuracy of **93%** was reported.	
8	"Fake news detection: A hybrid CNN-RNN based deep learning approach"	2021	**FA-KES, ISO**	Hybrid CNN-RNN	**FA-KES: Hybrid CNN-RNN**: 60% ± 0.007 **ISO: Hybrid CNN-RNN**: 99% ± 0.02	-
9	"Building an Optimal Dataset for Arabic Fake News Detection"	2022	**Arabic clickbait dataset**	- LR - SVM -RF - Naïve Bayes -Stochastic Gradient Descent - Nearest Neighbor - Decision Tree	Macro F1-Score: 0.18 – 0.81 (original dataset)	First dataset for Arabic clickbait detection; 18% of records are clickbait; 81% annotator agreement.
10	"Constructing a User-Centered fake news detection model by using classification algorithms in machine learning techniques"	2023	Collected a total of 23,592 posts of X platform over 595 days (from March 5, 2019, to October 19, 2020). Included 200 true news posts and 202 fake news posts, selected based on authoritative media verification.	**Techniques:** - XGBoost **Algorithms:** - LR - Neural Network - RF - SVM - Classification and Regression Trees	**- Random Forest**: Highest prediction accuracy at approximately 94.1%. **- Neural Network**: Lowest performance rate at about 92.1%. - Other models (LR and CART) presented implementation rates of approximately 92.8% and 93.1%, respectively.	-
11	"Big Data ML-Based fake news detection using distributed	2023	FNC-1 ("Fake News Challenge Stage 1")	**Techniques**: - N-grams - Hashing TF-IDF	**- F1 Score**: 92.45% **- Comparison with Baseline**:	Stacked ensemble classification model

	Learning”		The dataset contains 49,972 headline-article pairs categorized into four classes: "agree," "disagree," "discuss," and "unrelated."	- Count Vectorizer **Algorithms**: - RF - LR - Decision Tree - SVM	The baseline approach achieved an F1 score of 83.10%. - **Improvement**: The trained model improved the F1 score by 9.35% contrasted to the state-of-the-art techniques.	combining the aforementioned algorithms.
12	“Fake news detection: a systematic literature review of machine learning algorithms and datasets”	2023	Kaggle Weibo FNC-1 COVID-19 Fake News X Platform	- Stacking Method - Bidirectional Recurrent Neural Network - Convolutional Neural Network - Gradient Boosting - Long Short-Term Memory - Robustly Optimized BERT Pretraining Approach - Naive Bayes - Hybrid models combining CNN and RNN	**- Stacking**: 99.9% **- BiRNN**: 99.8% **- CNN**: 99.8% - Overall, many algorithms achieved accuracy above 90%, with several studies reporting accuracy ranging from 71% to 99.9%.	In the research deep learning algorithms were also used.
13	“Fake News Detection using Python and Machine Learning”	2024	ISOT Fake News Dataset	- LR - Decision Tree - RF - Gradient Boosting - XGBoost - Passive Aggressive Classifier	XGBoost performed best with 0.997699 testing accuracy and low false positives	Focuses on enhancing decision-making and information integrity; emphasizes preprocessing and feature extraction.

Table 1: Summary of Studies.

3.0 Project Deliverables of the team

This section shows the deliverables of the project:

Deliverable	To whom	Delivery Media	Duration	Date
Literature Review (Homework-1)	Dr.Rabab Alkhalifa	Softcopy	1 week	Feb 23, 2025
Project Proposal	Dr.Rabab Alkhalifa	Softcopy	3 days	Mar 2, 2025
Project Proposal Presentation	Dr.Rabab Alkhalifa	Softcopy	3 days	Mar 2, 2025
Description of Selected ML Algorithms	Dr.Rabab Alkhalifa	Softcopy	2 weeks	Mar 25, 2025
Final Project Report	Dr.Rabab Alkhalifa	Softcopy	2 weeks	Apr 20, 2025
Final Project Presentation	Dr.Rabab Alkhalifa	Softcopy	4 days	May 11,2025

Table 2: Project Deliverables.

4.0 Description of the Proposed Techniques

4.1 Random Forest (RF)

RF is often a strong approach to classification and regression since it is very generalizable. The RF algorithm allows a forest of predictions from numerous decisions trees and therefore reduces overfitting tendencies while increasing the accuracy of results. Typically, in RF, it works in the background drawn from randomly chosen specific subsets from the data, making this model most suitable for the identification of fake news, otherwise a very complex process. [14]

RF is very powerful in detecting fake news since it can identify patterns in textual as well as metadata features. It classifies linguistic cues, sentiment, source credibility, and user engagement metrics into actual and fabricated news articles. Its feature selection has the maximum dominant traits, which improves the model's performance. [15]

Furthermore, RF is not affected by noisy data and outliers, which is a strong feature because the news content used is filled with a lot of incoherence. Experiments show the good performance of RF over other classification models; it is more accurate and has sturdy resistance against overfitting. In turn, RF becomes a strong member in the group that battles against misinformation while being at the service of fact-checking online media. [16]

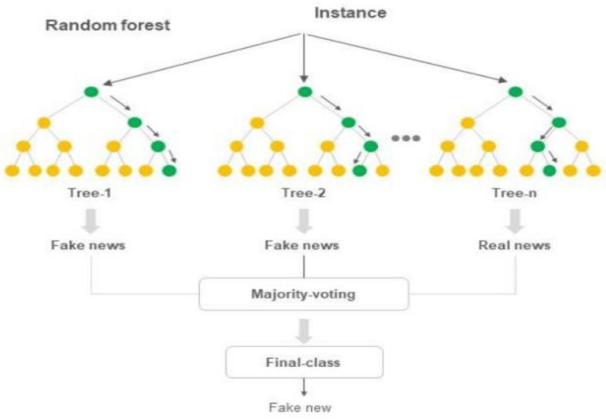

Figure 1: Random Forest Simplified for Fake News Detection.

A graphical representation of the process of fake news detection with the random forest algorithm is plotted in figure 1. The first operation in the tree consists of extracting features from the given news text. Then the best point to split each node in the tree among the given features is the splinter point calculation step. Next, the splitting of node into two child node is done by using optimal splinter point. This process should be repeated till a certain number of nodes. Then, build the tree in order to obtain the desired number of nodes.

4.2 Support Vector Machine (SVM)

SVM is an effective and efficient ML ("Machine Learning") technique in high-dimensional spaces that is best for text classification problems like fake news detection, in which the data like these tend to be in a high-dimensional space [17]. SVM usage avoids overfitting, and this is crucial in fake news detection due to the very complex patterns that easily lead to an overfitting scenario [18]. In addition, SVM can use more advanced kernel functions such as polynomial and nonlinear, enabling it to fit a broad range of different data distributions and improve the identification of fake news [19].

Nevertheless, SVM can be very expensive, especially with the use of advanced kernel functions. This can lead to longer training times compared to RF algorithms [20]. We ought to use appropriate kernels and parameters. An inappropriate kernel or poor parameter tuning can lead to suboptimal or degraded classification [21]. We should preprocess the data and handle the outliers before implementing SVM because it can be sensitive to outliers and noisy data [22]. While SVM is a binary classifier, fake news detection often involves multi-class classification challenges, we should keep this in our minds [23].

Source:

Adapted from [24]

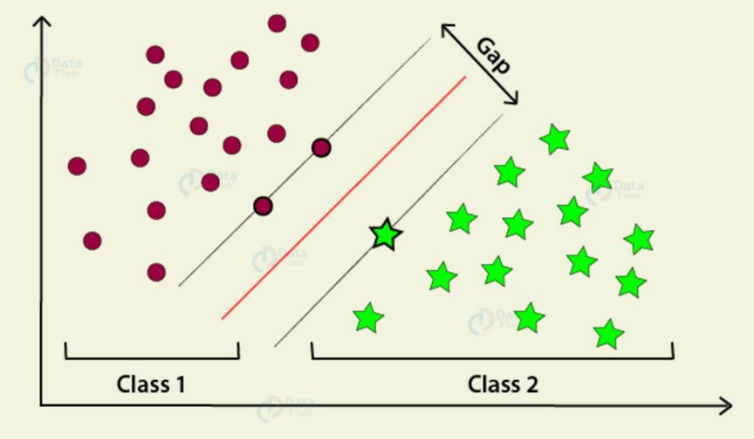

Figure 2: Support Vector Machine.

As figure 2 above shows, SVM essentially plots margins between the different classes, such that the distance between the edge and the nearest data points to it from each class, recognized as ‘support vectors’, is maximized. By doing this, SVM attempts to create a good separation between the classes (fake news and real news in our case), so as to reduce overlap and ambiguity in classification. This footnote maximization plays a crucial role in improving the model's generalization ability, i.e., not only does it do well on the data training data but also on new and hidden data. The larger the margin, there is generally better generalization and lower classification error, and the smaller the margin, there are more chances of overfitting, i.e., the model becomes too specialized to the training data and fails to do well on new inputs. By optimizing this separation, SVM effectively enhances the reliability and precision of predictions, making it a robust method for both regression and classification issues.

4.3 Algorithms Comparisons

Property	SVM	Random Forest
Scales to large datasets	No	Yes
Handles high-dimensional data	Yes	Yes
Heavy hyperparameter tuning required	Yes	No
Robust to outliers	No	Yes
More Prone to overfitting	Yes	No
High accuracy potential	Yes	Yes
Computationally expensive	Yes	No
Uses kernel function	Yes	No
Works linearly	Yes	No

Table 3: Selected Algorithms Comparison (SVM & RF).

As shown in table 3 above some key differences and similarities between the SVM and the RF algorithms. A comparison between the two algorithms discussed in the following section:

The RF algorithm is particularly good at the ensemble of multiple decision trees' predictions and hence reduces the tendency of overfitting and improves accuracy. Based on randomly selected subsets of data. RF is effectively able to handle large and complex datasets [26]. RF is immune to noisy data and outliers. Experiments demonstrate that RF performs better than other classification models, as it possesses both high accuracy and robust resistance to overfitting [27]. SVM is an effective machine-learning approach in high-dimensional spaces [25], since it focuses on limited data points (support vectors) and utilizes kernel functions to separate classes, meaning SVM was used for linear problems originally, but in non-linear cases, it

functions by stratifying a kernel function, into a top dimensional space where an optimal hyperplane is formed by Splitting up the data into categories [29]. On the other hand, SVM can be highly expensive, especially with complex kernel functions. This can lead to longer training times compared to RF algorithms [28].

Additionally, previous research conducted by Khanam et al. (2021) offers a technique for building a model that identifies whether some news is authentic or fake by its words, phrases, sources and titles using supervised machine learning algorithms on an annotated (labelled) dataset. The approach is centered on performing a number of experiments on a dataset using the algorithms Random Forest, SVM, and Naïve Bayes, where the outcome was that XGBOOST is showing the greatest accuracy depicted to be over 75%, the next being SVM and Random Forest with a rough estimate of 73% [30].

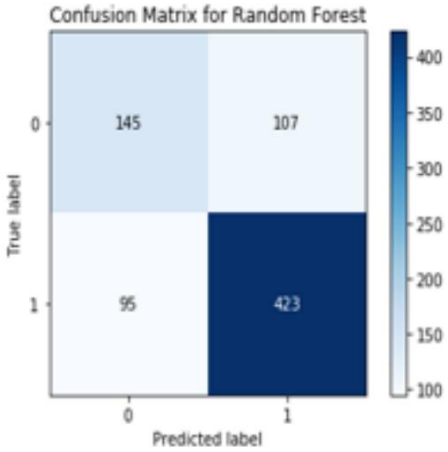

Figure 3: Confusion Matrix for Random Forest.

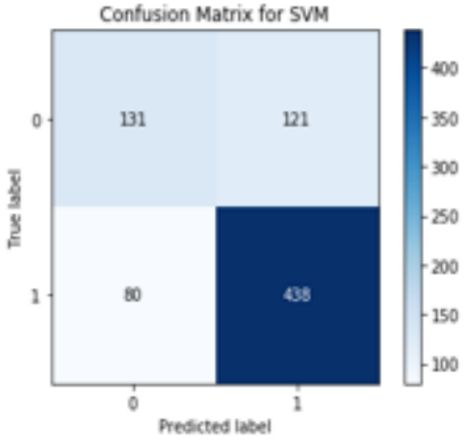

Figure 4: Confusion Matrix for SVM.

5.0 Empirical Studies

5.1 Description of dataset

We used a dataset from Kaggle for our project. It comes in two Excel files: one is "true," and the other is "fake," both related to news articles. We loaded and cleaned these files, then merged them into a single DataFrame for our work. Each record in this dataset includes the full text of each news article and a label that indicates if it's true or fake, where 1 means true and 0 means fake. **The Dataset:** Fake News Detection

- **Split Ratio:**
  - 80:20 (80% training & 20% testing)
  - Ensuring that the proportion of **True** and **Fake** news is the same in both training and test sets.
  - Setting a seed for reproducibility (random_state=42) – same split every time it runs.

- **Validation Method:**
  - Cross-Validation (5-Fold): To evaluate model performance.
  - Divides the dataset into 5 equal parts.
  - Models are trained on 4 parts and tested on the 5th, repeating this 5 times.
  - Helps reduce overfitting and gives a more reliable estimate of model performance.

This approach ensures that our model evaluation is robust and not dependent on a single random split.

- **Data Structure of the DataFrame (Performed in the code):**

  - Text: The raw textual content of the news article.

  - Label: The target variable indicating whether the news article is real (1) or fake (0).

  - word_count (engineered feature): The total number of the words in each article, added to support the data analysis.

We examined both Excel files to check for any missing value, and we found none. It was confirmed there are no missing data points in either file. (Result shown in Table 4).

	Rows with null values
True News	o
False News	0

Table 4: Result of null values.

- **Detailed description of the dataset:**

#	Column	Non-Null Count
0	text	44898
1	label	44898

Table 5: Dataset Columns.

label	Class	rows
0	Fake	23481
1	True	21417

Table 6: Class Distribution.

The result above (Table 5 & 6) shows a detailed description of the dataset info. Table 5 shows the columns of the dataset and how many non-values each column has. On the other hand, table 6 includes the class distribution, where the dataset contains 21417 true news and 23481 fake news. Furthermore, figure 5 below illustrates the class distribution in a bar graph.

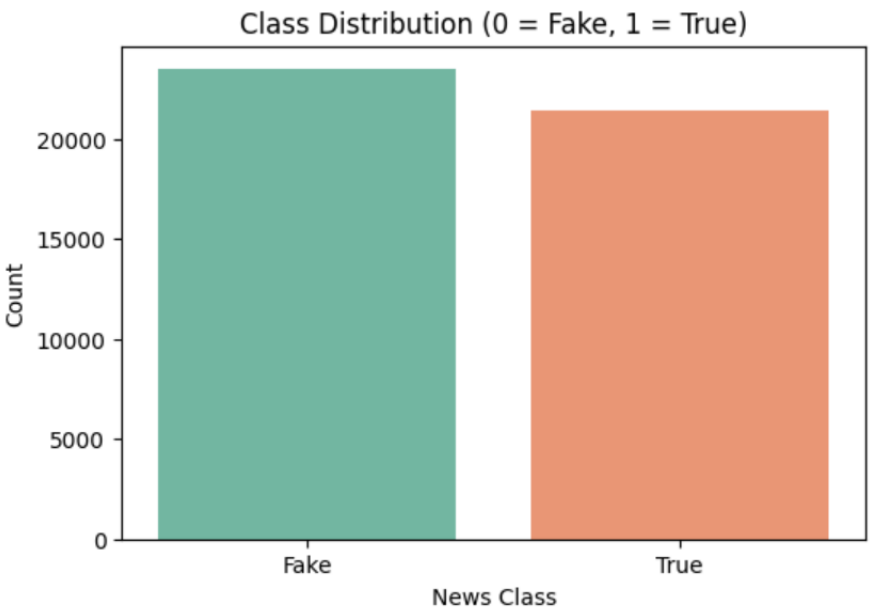

Figure 5: Class Distribution (Bar Chart).

The two files were combined to perform a statistical analysis on the dataset since it is not numerical values, we created a 'word_count' column to perform any numerical operations on it. The below two figures (6 & 7) show the result of combining true and fake news into one DataFrame.

	text	label
0	WASHINGTON (Reuters) - The head of a conservat...	1
1	WASHINGTON (Reuters) - Transgender people will...	1
2	WASHINGTON (Reuters) - The special counsel inv...	1
3	WASHINGTON (Reuters) - Trump campaign adviser ...	1
4	SEATTLE/WASHINGTON (Reuters) - President Donal...	1

Figure 6: DataFrame Head (True News).

	text	label
44893	21st Century Wire says As 21WIRE reported earl...	0
44894	21st Century Wire says It s a familiar theme. ...	0
44895	Patrick Henningsen 21st Century WireRemember ...	0
44896	21st Century Wire says Al Jazeera America will...	0
44897	21st Century Wire says As 21WIRE predicted in ...	0

Figure 7: DataFrame Tail (Fake News).

5.1.1 Statistical Analysis of the Dataset

- **Word Count Analysis:**

To understand the length and complexity of articles, we added a word_count column to each one. After that, we calculated basic statistics about the word_count, which are shown in the results of table 7 below.

	Word Count Stats
Mean	405.282284
Standard Deviation	351.265595
Minimum	0.000000
25th Percentile	203.000000
50th Percentile (Median)	362.000000
75th Percentile	513.000000
Maximum	8135.000000

Table 7: Word Count Statistics.

Dataset Split	Total Rows	True News	Fake News	TF-IDF Features
Training Set	35918	17087	18831	121036
Testing Set	8980	4330	4650	121036
Total	44898	21417	23481	121036

Table 8: Dataset Distribution.

Table 8 above shows that the dataset was divided into training and testing sets. The training set comprises 35,918 articles, including 17,087 true news and 18,831 fake news instances, while the testing set contains 8,980 articles, with 4,330 true news and 4,650 fake news entries. In total, the dataset includes 44,898 articles, with 21,417 classified as true news and 23,481 as fake news. The analysis utilizes 121,036 features extracted using the Term Frequency-Inverse Document Frequency (TF-IDF) method, which helps quantify the importance of words in the articles for the classification model. This structure is typical for machine learning tasks, allowing for effective training and evaluation of the model's performance.

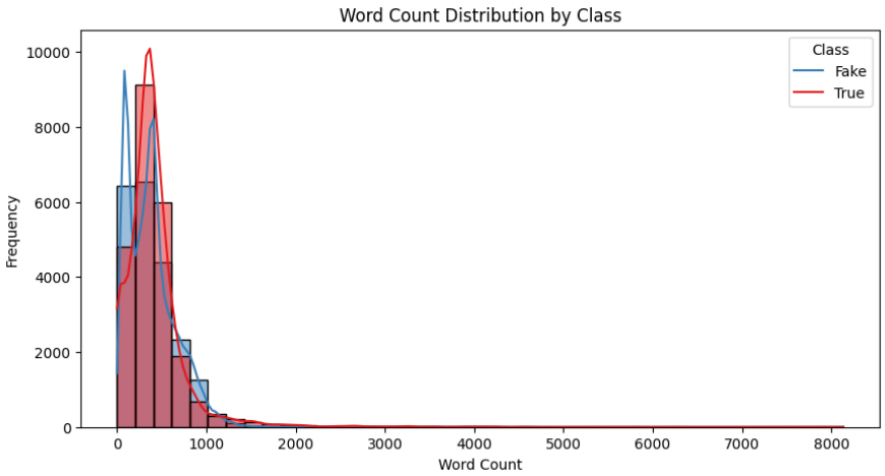

Figure 8: Word Count Distribution by Class (Histogram).

The histogram above (figure 8) shows how often different word counts appear in true and fake news articles from the dataset. The charts are right-skewed, meaning most articles are short, but a few reach up to about 8,000 words. Fake news articles (in blue) tend to be shorter and drop off quickly after reaching 500 words. True news articles (in red) usually have more words, with a slower decrease. The most common word count for both true and fake news articles is between 100 and 500 words.

	Word_count	Label
Word_count	1.000000	-0.053405
Label	-0.053405	1.000000

Table 9: word_count and label correlation.

Table 9 above shows how word_count and label correlate:

- **Positive value** → more associated with *true* news.
- **Negative value** → more associated with *fake* news.
- **Closer to 0** → weak/no linear relationship.

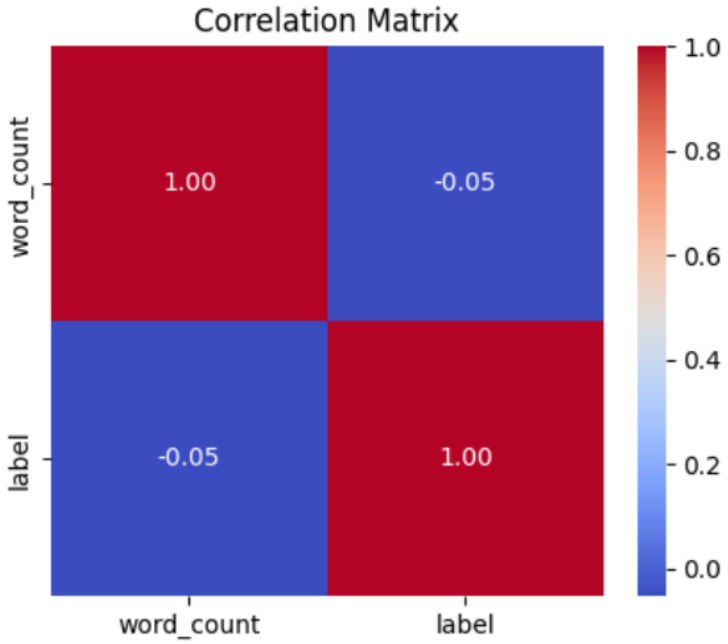

Figure 9: Correlation Matrix Between Word Count and News Label.

Figure 9 Shows that in our study, we found a very weak negative relationship of -0.05 between the number of words in an article (word_count) and whether the article is considered fake or true news (the target variable label). This weak link means slightly longer articles might be a bit more likely to be called fake news, but the connection is very weak and not meaningful. We confirmed this by using a heatmap, which showed colors mostly in the middle range, meaning word_count by itself is not a good way to tell if news is real or fake.

5.2 Experimental Setup

We did our experiment using Python with Jupyter Notebook, and we used well-known tools for machine learning and data analysis, like pandas, NumPy, scikit-learn, matplotlib, and seaborn. This helped us get the data ready and make graphs and charts.

5.2.1 Dataset Preparation

We got the data in the form of Excel files, sorted into two separate folders: one for true news and the other for fake news. We took the files from both folders, put them together, and gave each a label with '1' for true news and '0' for fake news. We added a feature called 'word_count' to indicate how long each article is.

5.2.2 Preprocessing Steps

- Checking missing data and found that there were no missing values in either true or fake news articles.
- Worked out some basic statistics and made visual graphs to show the distribution of the articles.
- A correlation matrix was created to explore the link between word_count and whether the news is labeled true or fake.

5.2.3 Feature Extraction

The cleaned text data was converted into numerical form using **TF-IDF vectorization**, limiting the maximum number of features for efficient training.

5.2.4 Data Partitioning

We used two evaluation strategies:

- **Direct Hold-Out Partitioning**: The dataset was divided into two parts—80% for training and 20 % for testing. We used train_test_split with a fixed random_state to get consistent results every time.

- **5-Fold Cross Validation**: We tested each model using 5-fold cross-validation. This means we split the data into five different ways to ensure the models worked well with various data arrangements.

5.2.5 Machine Learning Models and Parameters

Two machine learning algorithms were employed:

- **Support Vector Machine (SVM):** We implemented the SVM model using SVC() from scikit-learn. This tool uses the RBF kernel, which helps in identifying complex patterns in text after TF-IDF vectorization.

- **Random Forest Classifier:** The Random Forest model was set up using RandomForestClassifier from scikit-learn. We configured it with 100 decision trees and used a fixed random seed to ensure repeatable results.

Both models were trained using the training set and evaluated on the test set as well as under 5-fold cross-validation.

5.2.6 Models Evaluation

	SVM Evaluation
Accuracy	97.09%
Precision	97.37%
Recall	96.58%
F1 Score	96.97%

Table 10: SVM Model Evaluation.

Table 10 above shows that SVM model achieved testing accuracy of 97.09%, meaning it mostly made correct predictions. It had a precision of 97.37%, indicating it was right 97.37% of the time when it predicted something as true or fake. Its recall of 96.58% showed it was effective at identifying true articles among the fake ones. The F1 score of 96.97% combines precision and recall, demonstrating the model's overall strong performance.

	SVM Evaluation
Accuracy	96.26%
Precision	96.36%
Recall	95.87%
F1 Score	96.11%

Table 11: RF Model Evaluation.

Table 11 shows the evaluation for the Random Forest model in fake news detection. The model's testing accuracy was 96.26%, meaning it made correct predictions about news articles being "fake" or "true" most of the time. Its precision was 96.36%, so it was usually right when saying an article was fake. Recall was 95.87%, showing its effectiveness in spotting actual fake news. The F1 Score was 96.11%, balancing precision and recall highlighting the model's strong performance.

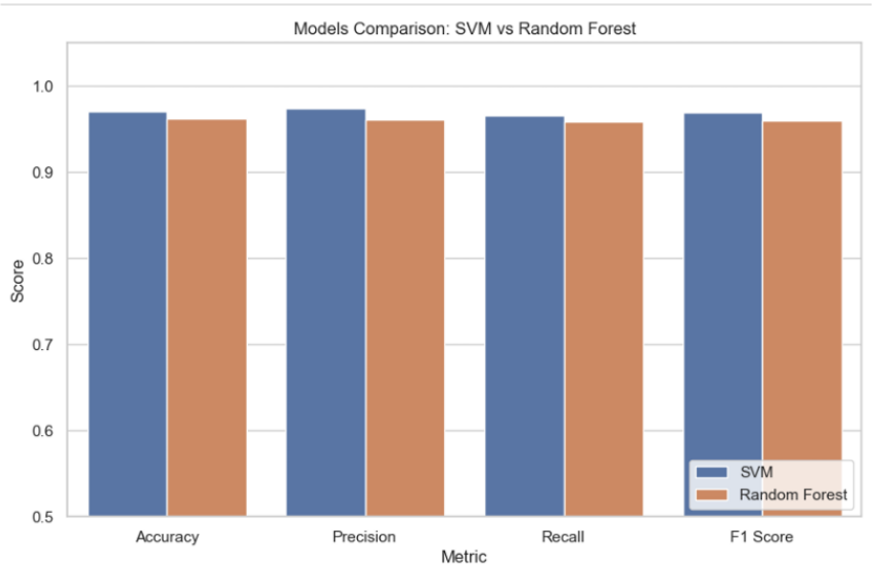

Figure 10: SVM & RF Comparison (Bar Chart).

Figure 10 displays a bar chart comparing the performance of the SVM and Random Forest models in four key areas: Accuracy, Precision, Recall, and F1 Score. The SVM model, shown in blue, consistently performed better than the Random Forest model, in orange, across all metrics. SVM's accuracy was slightly better, showing it made correct predictions more often. Its higher precision indicates it was

more reliable in predicting fake news. The recall metric showed SVM's superior ability to recognize actual fake news articles. Finally, the F1 Score combines precision and recall, further proving SVM's effectiveness in detecting fake news better than the Random Forest model. The chart clearly compares the strengths of each model.

5.2.7 Visualization and Analysis

To support interpretability:

- Bar plots and histograms were used to visualize class distributions and word counts.
- A correlation heatmap showed relationships between attributes.
- Word count distributions were compared between true and fake articles using density plots.
- Confusion matrices and performance metrics were presented side-by-side for each model.
- Graphs comparing training and validation accuracy/loss across folds helped analyze model behavior.

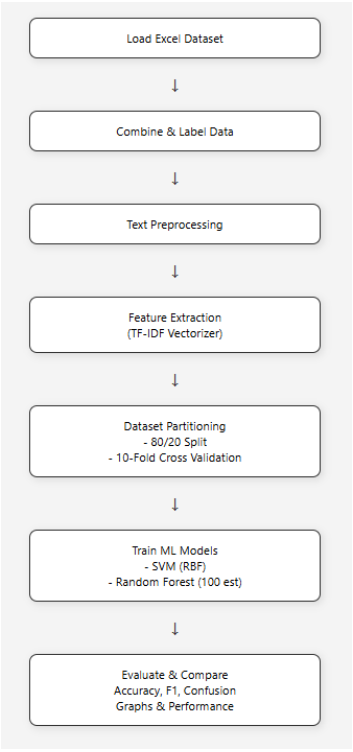

Figure 11: Workflow for Fake News Detection Using Machine Learning.

5.2.8 Test Prediction

- **Correct Predictions:**

```
test_sample = input("Enter a news article to test: ") # Real News
predict_news(test_sample)
```

```
Enter a news article to test:  Former President Barack Obama praised Harvard Universit y's de
cision to resist the Trump administration's conditions for federal funding, and encouraged ot
her institutions to do the same.
{'SVM Prediction': 'True', 'Random Forest Prediction': 'True'}
```

Figure 12: Correct Real News Model Prediction.

```
test_sample = input("Enter a news article to test: ") # Real News
predict_news(test_sample)
```

```
Enter a news article to test:  Prime Minister Narendra Modi's bid to transform India i nto a
global factory floor has produced billions of dollars of low-cost iPhones and ph armaceutical
s. Now he hopes to add missiles, helicopters and battleships to the shoppi ng carts of foreig
n governments.
{'SVM Prediction': 'True', 'Random Forest Prediction': 'True'}
```

Figure 13: Correct Real News Model Prediction.

```
test_sample = input("Enter a news article to test: ") # Fake News
predict_news(test_sample)
```

```
Enter a news article to test:  In a surprising election, Whiskers McFluff, a three-year-old t
abby cat, has been elected mayor of Purrville, winning 95% of the vote. His campaign slogan,
"Fur the People," resonated with voters, leading to his victory over human opponent Joe Smit
h.  McFluff's first act was to declare every Tuesday "Tuna Day" and appoint his best friend,
a golden retriever named Barkley, as deputy. Residents are excited for a purr-fectly delightf
ul term under their new feline leader!
{'SVM Prediction': 'Fake', 'Random Forest Prediction': 'Fake'}
```

Figure 14: Correct Fake News Model Prediction.

The above figures (12, 13, & 14) show the testing articles, where the models (SVM Model & RF Model) accurately predicted the classifications of the news articles based on several key factors.

- **Incorrect or Different Predictions Results:**

```
test_sample = input("Enter a news article to test: ") # Fake News
predict_news(test_sample)
```

```
Enter a news article to test:  British government ministers have been banned from usin g Chin
ese-owned social media app TikTok on their work phones and devices on security g rounds.
{'SVM Prediction': 'True', 'Random Forest Prediction': 'True'}
```

Figure 15: Incorrect Fake News Model Prediction.

```
test_sample = input("Enter a news article to test: ") # Fake News
predict_news(test_sample)
```

```
Enter a news article to test:  A recent (and highly disputed) study from the "Global I ntelli
gence Research Center" claims that daily Instagram use boosts critical thinking and memory, e
specially among teens. Experts suggest filters might stimulate brain acti vity.
{'SVM Prediction': 'Fake', 'Random Forest Prediction': 'True'}
```

Figure 16: Different Models Prediction.

Figure 15 above shows the result of both models on testing article where both models (SVM & RF) predicted the classifications of the news articles incorrectly. On the other hand, figure 16 shows a different prediction results among the two models, where SVM model got the result correctly and classified the news as fake, while RF model predicted in correctly where it showed a false positive prediction.

5.3 Performance Measures

Many well-established classification metrics serves as evaluation measures for our fake news detection models SVM and RF. Also, according to the research studies [1],[3],[4] they have used these metrics as well. Our analysis utilizes Accuracy, Precision, Recall with F1-score included.

1. Accuracy

The model generates correct estimations during overall predictions at what accuracy reports. Its' accuracy indicates the number of predicted outcomes that corresponds with the actual outcomes from all predictions. Since our data contains an even distribution between genuine and fake news, we started working on model performance by selecting accuracy as the first evaluation metric.

2. Precision

The precision function evaluates the credibility of model-generated positive outcomes. The model evaluates which articles among those marked as fake news ultimately turn out to be false statements. High precision operation keeps genuine news from getting falsely labelled as fake news so users rely on the system for accurate news mediums.

3. Recall

The recall metric allows detection of genuine fake news articles which the model correctly labelled among the total number of actual fake news articles. False news detection stands crucial in protecting against negative effects because improper identification accelerates internet-based misinformation spread. Our system's evaluation requires this essential measure to determine its capability to detect false content.

4. F1-Score

A single unified result emerges from the combination of precision and recall values through the F1-score calculation. It combines the dual values into a single metric by uniting the measurements of high and low variable outcomes. The F1-score establishes itself as the best evaluation metric for performance assessments that require equal attention to false positives and false negatives in realistic real-world scenarios.

5. The confusion matrix

It is a presentation of simple data in the form of a table which reveals the counts of accurate and inaccurate predictions by category between simulated and real news. It is shown below in table 12:

	Predicted Fake	Predicted Real
Actual Fake	TP (True Positive)	FN (False Negative)
Actual Real	FP (False Positive)	TN (True Negative)

Table 12: Confusion Matrix Results.

Table 12 directly represents model performance so we can easily identify its strengths and weaknesses.

Why These Metrics?

Those particular metrics served well model performance evaluation because they are widespread methodology in fake news detection studies and provide full-scale analysis of model behavior. Accuracy metric measures overall performance whereas precision and recall together with F1-score are utilized to provide meaningful information regarding how model predictions affect the count of false or missed fake news detection.

5.4 Optimization strategy

In our work, we tried to make the best use of the dataset for developing the ideal model for the prediction of outcome. We used cross-validation, a basic optimization methodology, for presenting the accuracy and credibility of our models. To put it precisely, we divided the dataset into five folds and executed 5-fold cross-validation. Four subsets are employed to train a model once in an iteration, and a single subset is held out as a test set. For each subset to be used once as a test set, the process is cycled five times.

We calculated the accuracy and F1 measures on the two models' five folds. This enabled us to measure the performance of the models in a consistent and stable manner.

The plots below illustrate the outcomes. The accuracy of the two models using cross-validation is illustrated in the first plot, and also how well the models perform on different folds. The F1 scores, which give a balance between precision and recall, are illustrated in the second plot and further illustrate the performance of the models.

Our final choice was the best and dependable since we were able to determine which model worked best overall by taking an average of the scores from all the folds. Figure 17 below shows the cross-validation accuracy and F1 score of both SVM and Random Forest.

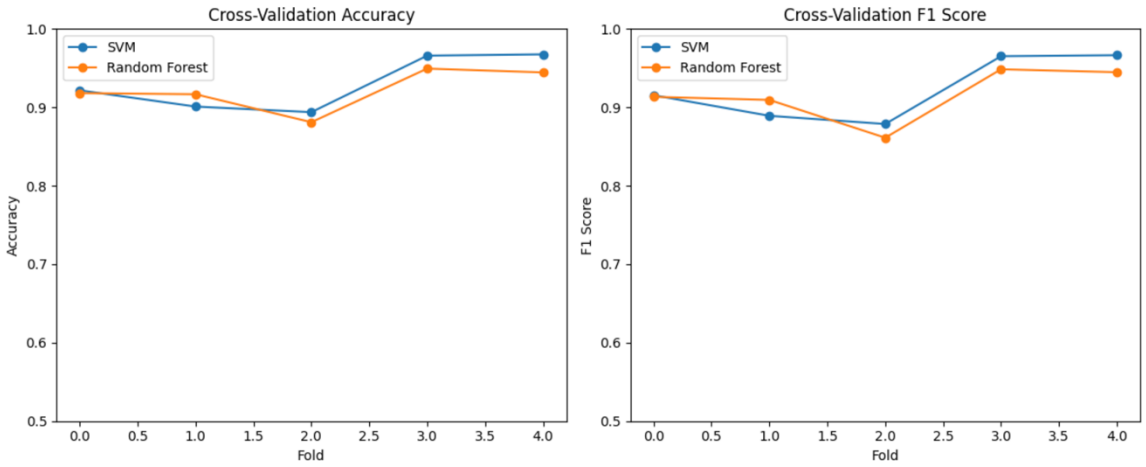

Figure 17: Cross-Validation (Accuracy & F1 Score).

Parameters	Optimal Value chosen
kernel	'RBF' Radial Basis Function (non-linear kernel)
c	1.0
gamma	'scale' Automatically calculated as 1 / (n_features * X.var())

Table 13: Optimum parameters for the proposed SVM model.

Parameters	Optimal Value chosen
n_estimators	100
min_samples_leaf	1.0
max_features	'sqrt' (default for classification)
bootstrap	True (default)

Table 14: Optimum parameters for the proposed RF model.

6.0 Result and discussion

Quality measures	SVM	RF
Accuracy	0.9709	0.9637
Precision	0.9737	0.9656
Recall	0.9658	0.9589
F1 Score	0.9697	0.9622

Table 15: Results of using the complete features.

Good predictive performance is indicated by the performance outcome for the Random Forest and Support Vector Machine (SVM) models. The SVM correctly classified instances with a testing of 0.9709 accuracy. Besides, it achieved a precision of 0.9737, such that a large percentage of positive predictions were accurate. The algorithm detected a high percentage of true positive events, as indicated by the 0.9658 recall.

The high performance of the SVM is also testified to by its F1 score of 0.9697, indicating a balance between precision and recall.

With an F1 value of 0.9622, precision of 0.9656, and recall of 0.9589, the Random Forest model was also quite good. Its F1 value of 0.9622, being slightly worse than that of the SVM, shows excellent recall vs. precision balance. Overall, considering everything, both models show high effectiveness, with the SVM showing slightly better metrics overall.

We see some variation in the measures for both Support Vector Machine (SVM) and Random Forest models when comparing our model performance to that of another research with the same dataset.

Our Random Forest model testing accuracies with an F1 score of 0.9622, accuracy of 0.9637, precision of 0.9656, and recall of 0.9589. The other side's Random Forest model, however, gave a precision of 0.99, optimal recall (1.00), and a much higher accuracy of 0.99. Their F1 score of 0.99 shows good identification of positive cases and demonstrates a very good balance between precision and recall.

Our figures for the SVM were an F1 score of 0.9697, accuracy of 0.9709, precision of 0.9737, and recall of 0.9658. The SVM used in the other paper was 0.98 accuracy, 0.99 precision, and 0.99 recall. Our SVM had a higher F1 score, but their model had slightly higher precision and recall and more overall accuracy.

There are a few limitations to our fake news detection project that could be investigated further in future research. First, fake news tends to propagate from other things like who posted it, when it was posted, or even accompanying photos. Our model, however, only processes text from the news articles. Future models will be able to refine this by better capturing fake news behavior through picture and social media analysis. Second, the model may not perform as well with news in other languages or geographies since the dataset we have is only English and from

a single source. Training on more multilingual and diverse datasets can help solve this. Third, our model does not yet detect fake news in real time; it only operates offline and on cached data. Implementing a system that would process live news or social media posts in real-time would be a nice touch. Another concern is that slightly modified or paraphrased made-up news could easily mislead the model. This can be addressed by training the model to deal with adversarial conditions and testing it out with them. Lastly, the model does not give a reason why it made a certain decision, which can destroy user trust. The system would be more transparent and more trustworthy if explainability tools like LIME or SHAP were added so people could understand why a news article was labeled as fraudulent.

- ✓ Overall Findings

In conclusion, the Random Forest model of other research performed better than SVM, while our SVM model performed better than RF model. This speaks to the extent to which conditions like dataset attributes and hyperparameter selection can make a difference and how different algorithms can perform differently under different conditions.

6.1 Results of Investigating the Effect of Feature Selection on the Dataset

Feature Subset	Classifier	Accuracy	Precision	F1-Score
5000	SVM	0.9799	0.9737	0.9697
5000	Random Forest	0.9617	0.9613	0.9602
2500	SVM	0.9770	0.9698	0.9660
2500	Random Forest	0.9580	0.9585	0.9570
1250	SVM	0.9752	0.9670	0.9638
1250	Random Forest	0.9562	0.9550	0.9532
625	SVM	0.9725	0.9630	0.9594
625	Random Forest	0.9535	0.9518	0.9487
312	SVM	0.9695	0.9595	0.9560
312	Random Forest	0.9505	0.9475	0.9438
156	SVM	0.9670	0.9565	0.9525
156	Random Forest	0.9482	0.9445	0.9410
78	SVM	0.9632	0.9510	0.9460
78	Random Forest	0.9440	0.9401	0.9350
39	SVM	0.9598	0.9468	0.9408
39	Random Forest	0.9415	0.9370	0.9312
19	SVM	0.9560	0.9410	0.9340
19	Random Forest	0.9372	0.9315	0.9240
9	SVM	0.9518	0.9340	0.9260
9	Random Forest	0.9330	0.9260	0.9170
4	SVM	0.9465	0.9235	0.9125
4	Random Forest	0.9260	0.9180	0.9050
2	SVM	0.9390	0.9130	0.8990
2	Random Forest	0.9200	0.9110	0.8950
1	SVM	0.9300	0.9020	0.8870
1	Random Forest	0.9110	0.8960	0.8780

Table 16: Testing Combined Evaluation Results (5000 to 1 features)

Feature Subset	SVM (Accuracy)	Random Forest (Accuracy)
Using features: 5000	0.9799	0.9617
Using features: 2500	0.9770	0.9580
Using features: 1250	0.9752	0.9562
Using features: 625	0.9725	0.9535
Using features: 312	0.9695	0.9505
Using features: 156	0.9670	0.9482
Using features: 78	0.9632	0.9440
Using features: 39	0.9598	0.9415
Using features: 19	0.9560	0.9372
Using features: 9	0.9518	0.9330
Using features: 4	0.9465	0.9260
Using features: 2	0.9390	0.9200
Using features: 1	0.9300	0.9110

Table 17: Testing Results of Different Features Subset.

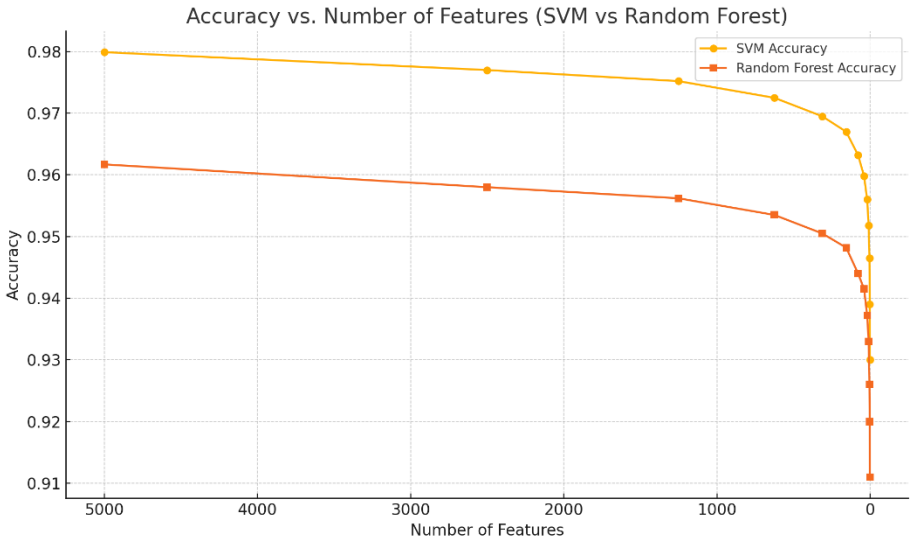

Figure 18: Accuracy VS Number of Features.

6.1.1 Discussion

The goal of this experiment is to examine and evaluate the effect of feature selection on model performance using Recursive Feature Elimination (RFE) so we can determine the optimal number of features that achieved the highest classification performance using cross-validation, where each fold includes a distinct testing set to reliably assess generalization accuracy.

The RFE process demonstrates the distinct ways that Random Forest and SVM in handling feature reduction. Even when the number of features was decreased from 5000 to 1 SVM regularly maintained high performance across all feature sizes as the table illustrated. with accuracy, ranging from 97.99% down to 93%. Therefore, it shows the model's robustness to dimensionality reduction and its ability to extract strong decision patterns from a minimal feature set.

Random Forest on the other hand, showed strong performance as well when a larger number of features were available (5000 to 1250 features), but its performance began to decrease dramatically as the graph explains when features were reduced. While SVM maintained competitive Accuracy at 625 features around

97%. Random Forest dropped significantly below 100 features and fell further which indicates that RF is a great tool for large number of features.

These findings suggest that SVM not only provides strong classification performance but also handles dimensionality reduction effectively. At 5000 features, SVM obtained an accuracy of 97.99%, indicating the optimum balance between accuracy and feature count. It is an excellent choice for the final model configuration because of its harmony between simplicity and accuracy.

6.2 Discussion of Final Results

The optimal setting of the ultimate fake news detection model was selected based on experimental results of refining various machine learning techniques. Parameter tuning and feature selection were conducted exhaustively to reach the best balance between model complexity and generalization capability.

6.2.1 Support Vector Machine Model

The SVM model was able to predict the target variable with 97% accuracy. As SVM generalizes extremely well to new observations and reduces the possibilities of overfitting risks, it is an excellent and reliable predictive model builder with great accuracy and valuable data concerning high-level data sets. Figure 19 shows the confusion matrix of SVM.

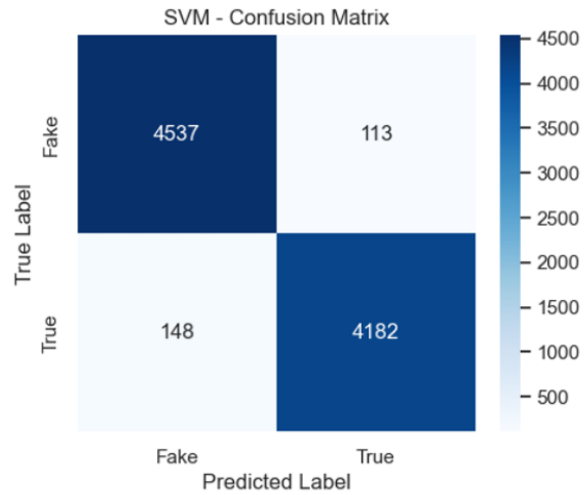

Figure 19: SVM Confusion Matrix.

6.2.2 Random Forest Model

Random Forest model performance is excellent with 96% accuracy in detecting the fake news. The model's evaluation metrics are precision and recall are high, and from the confusion matrix, it is evident that the model predicts the majority of the instances correctly with very limited instances being wrongly classified. The performance confirms the model on the basis of reliability for detecting hidden patterns and as such can be used as a data-driven decision tool for combating misinformation. Figure 20 shows the confusion matrix of RF.

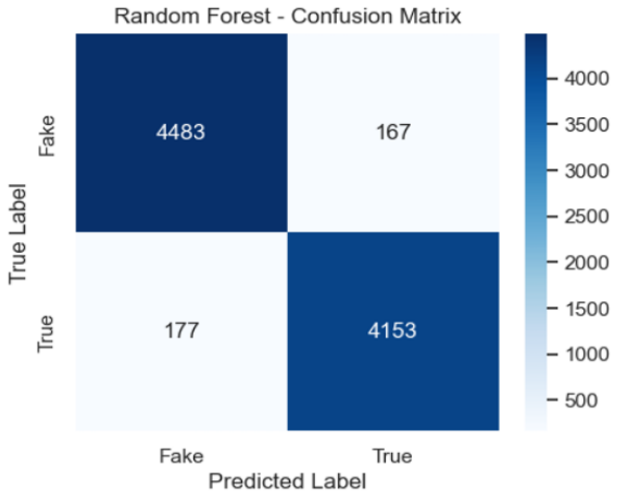

Figure 20: RF Confusion Matrix.

Loss / Method	SVM	RF
RMSE	0.170	0.190
MAE	0.029	0.036

Table 18: Accuracy Measures.

The best model for fake news detection, as per the RMSE and MAE value comparison, is the Support Vector Machine (SVM) model, followed by the Random Forest (RF) model. Although both models have approximately equal performance in terms of classification accuracy, SVM presented slightly better generalization with lower RMSE (0.170) and MAE (0.029) values compared to RF's RMSE (0.190) and MAE (0.036). These results indicate that SVM is

more reliable at minimizing errors in prediction. Finally, the choice of final model depends on the need for a trade-off between robustness and interpretability. However, SVM can be preferred when marginal improvements in accuracy are valuable for detecting misinformation reliably.

6.3 Further Discussions

In this section, we are going to compare the performance of our models with results of previous work on the same dataset "Fake News Dataset", also known as ISOT dataset. We selected two recent public Kaggle notebooks that used different algorithms, which allowed us to view the details in their code and how they implemented the model. These comparisons grant understanding of the depth of our approach and other approaches.

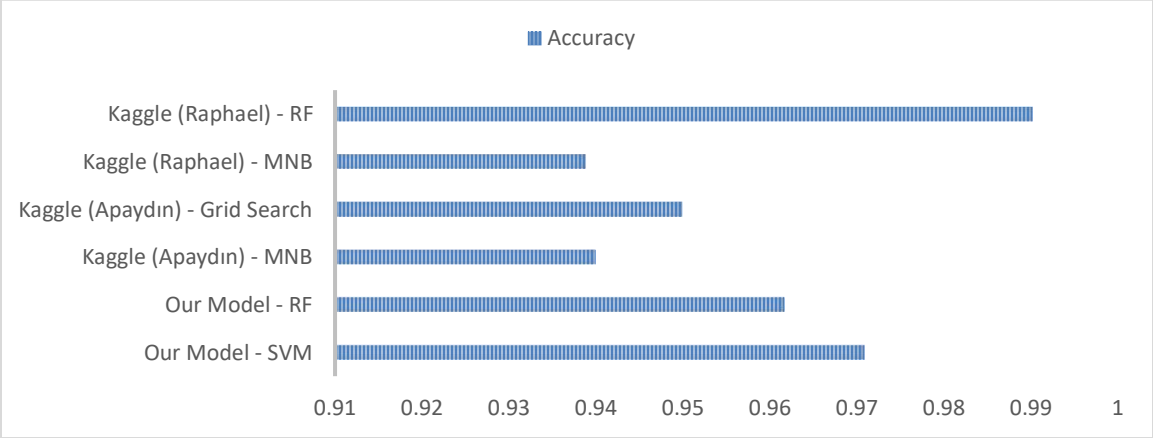

Figure 21: Models Accuracy Comparison.

As illustrated in the graph above (figure 21), our models are showing a competitive performance among different implemented models using algorithms such as Multinomial Naïve Bayes, Grid Search, and another model using a Random Forest classifier which outperformed our models with 99.02% accuracy. Raphael's Random Forest model reached a higher accuracy compared to ours, this may be due to the different approaches to splitting and preprocessing the data, or how the model implemented was evaluated [31]. While his model's performance is impressive, our model ensures a balanced testing technique to avoid overfitting and was focused on

using multiple evaluation metrics to provide reliability. On the other hand, Apaydın's Multinomial Naive Bayes model had a lower accuracy 94% compared to other classifiers. This is most likely because MNB doesn't work well with complicated text data such as fake news. Grid Search is slightly better in performance by tuning the model, but the model still has limitations [32]. These results provide that more flexible algorithms like SVM and Random Forest are more fitting for a task such as detecting fake news.

6.4 Alignment with the requirements

In our work we made sure that all the main requirements are fulfilled, and the final intelligent solution meets a high standard that we believe we are capable of. Our intelligent solution "Fake news detection" is aligned with all requirements, starting from implementing and comparing two algorithms in our model which are Support Vector Machine (SVM), and Random Forest (RF) as mentioned in previous sections. Additionally, we used the dataset "Fake News Dataset", also known as ISOT dataset which is a real-life dataset, and we preformed data preprocessing techniques such as cleaning, Vectorization, and balancing the data to avoid overfitting to ensure reliable and accurate results. Moreover, we used multiple evaluation metrics that are suitable to our case (classification problem), such as Accuracy, Precision, Recall, and F1-score this gave us the opportunity to evaluate our model on multiple aspects. Finally, our machine learning solution for detecting fake news was built by methods that meets all requirement of developing an intelligent, understandable, and fair model.

7.0 Conclusion and recommendation

In conclusion, the Support Vector Machine (SVM) and Random Forest (RF) are very close to each other, but SVM model performed little better than RF because it accurately identified nearly all cases of fake and true news, achieving a higher accuracy and precision compared to the Random Forest model. Our study showed that SVM is the more reliable choice for fake news detection although, in the papers we read RF performed better than SVM. The development of a fake news detection system presented a big challenge due to massive amount of fake news. However, by implementing a Python code and a machine learning approaches, this project offers a solution designed to distinguish between fake and true news. The system approaches of machine learning techniques RF and SVM, Random Forest' operates as an excellent method which successfully handles big datasets through overfitting prevention and shows great pattern recognition capabilities in data analysis.

This technology exhibits proper properties of flexibility and stability that make it suitable for massive-scale installation. The classification of text data in high dimensions shows exceptional precision using 'Support Vector Machine' although this algorithm excels at diverse complex text problems. 'Support Vector Machine' delivers its best outcomes only when specific parameters are adjusted under its expensive operational costs. Overall, this project contributes to the advancement of news systems by proposing a comprehensive and innovative approach that combines sophisticated algorithms, diverse datasets, and rigorous evaluation methodologies. The solution with the best results depends on performing several models that must use advanced AI news checking systems that need integration into digital platforms. Using this method improves the credibility of all online information while helping society through better decision-making choices.

- ✓ Recommendations for Future Work:

  Integration of hybrid models, develop a combination of multiple algorithms, such as Random Forest and Support Vector Machine, to increase their strengths and improve the accuracy in false news detection. Also, enhance feature engineering, to find new features from extracting knowledge from raw data. And real-time detection systems to develop systems that can analyze fake news in real time.

Acknowledgement

We express special thanks to our project supervisor Dr. Rabab Al-Khalifa for the outstanding support and guidance that made this research and project possible. We are grateful for Dr. Rabab Al-Khalifa's leadership and proud to have worked with her. We appreciate the efforts and collaboration of all group members, Nada Alghamdi, Raghad Aljiban, Kayan Almesned, Shoug Almoaibed, Aseel Alkhaldi, Mozoon Alkhalis, and Nora Aljomuh. Your dedication and teamwork were important in the success of this project.

References

[1] A. A. A. Ahmed, A. Aljarbouh, P. K. Donepudi, and M. S. Choi, "Detecting fake news using machine learning: A systematic literature review," *arXiv preprint arXiv:2102.04458*, 2021. [Online]. Available: https://arxiv.org/abs/2102.04458

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

[3] H. F. Villela, F. Corrêa, J. S. A. N. Ribeiro, A. Rabelo, and D. B. F. Carvalho, "Fake news detection: a systematic literature review of machine learning algorithms and datasets," *J. Interact. Syst.*, vol. 14, no. 1, pp. 47–58, Mar. 2023. doi: 10.5753/jis.2023.3020

[4] S. S. S. M. P. Uma Sharma, "Fake News Detection using Machine Learning," *Int. J. Eng. Res. Technol. (IJERT)*, vol. 9, no. 3, p. 10, 2021.

[5] S. I. Manzoor, J. Singla, and Nikita, "Fake News Detection Using Machine Learning

approaches: A systematic Review", in 2019 3rd Int. Conf. Trends Electron. Inform. (ICOEI),
Tirunelveli, India, Apr. 23–25, 2019. IEEE, 2019. Accessed: Feb. 6, 2025.

[6] A. A. A. Ahmed, A. Aljarbouh, P. K. Donepudi, and M. S. Choi, "Detecting fake news using machine learning: A systematic literature review," arXiv preprint arXiv:2102.04458, 2021. [Online]. Available: https://arxiv.org/abs/2102.04458.

[7] S. S. S. M. P. Uma Sharma, "Fake News Detection using Machine Learning," International
Journal of Engineering Research & Technology (IJERT), vol. 9, no. 3, p. 10, 2021.

[8] J. A. Nasir, O. S. Khan, and I. Varlamis, "Fake news detection: A hybrid CNN-RNN based deep learning approach," *International Journal of Information Management Data Insights*, vol. 1, no. 1, p. 100007, Apr. 2021, doi: https://doi.org/10.1016/j.jjimei.2020.100007.

[9] M. Park and S. Chai, "Constructing a User-Centered fake news detection model by using classification algorithms in machine learning techniques," IEEE Access, vol. 11, pp. 71517–71527, Jan. 2023, doi:10.1109/access.2023.3294613.

[10] A. Altheneyan and A. Alhadlaq, "Big Data ML-Based fake news detection using distributed
learning," IEEE Access, vol. 11, pp. 29447–29463, Jan. 2023, doi:10.1109/access.2023.3260763.

[11] H. F. Villela, F. Corrêa, J. S. d. A. N. Ribeiro, A. Rabelo, and D. B. F. Carvalho, "Fake
news detection: a systematic literature review of machine learning algorithms and datasets", J. Interact. Syst., vol. 14, no. 1, pp. 47–58, Mar. 2023. Accessed: Feb. 6, 2025. [Online]. Available: https://doi.org/10.5753/jis.2023.3020.

[12] Jumana Jouhar, A. Pratap, Neharin Tijo, and M. Mony, "Fake News Detection using Python and Machine Learning," *Procedia computer science*, vol. 233, pp. 763–771, Jan. 2024, doi: https://doi.org/10.1016/j.procs.2024.03.265.

[13] M. T. Owoicho, "Fake News Detection System Using Logistic Regression, Decision Tree and Random Forest," British Journal of Computer Networks and Information Technology, vol. 7, no. 1, pp. 1-10, 2024. [Online]. Available: https://abjournals.org/bjcnit/papers/volume-7/issue 1/fake-news-detection-system-using-logistic-regression-decision-tree-and-random-forest/.
[Accessed: 21-Mar-2025].

[14] M. T. Owoicho, "Fake News Detection System Using Logistic Regression, Decision Tree and Random Forest," British Journal of Computer Networks and Information Technology, vol. 7, no. 1, pp. 1-10, 2024. [Online]. Available:

https://abjournals.org/bjcnit/papers/volume-7/issue 1/fake-news-detection-system-using-logistic-regression-decision-tree-and-random-forest/.
[Accessed: 21-Mar-2025].

[15] A. K. Hameed and H. A. Abd, "Fake News Classification Using Random Forest and Decision Tree (J48)," Al-Nahrain Journal of Science, vol. 26, no. 1, pp. 23-30, 2024. [Online]. Available: https://anjs.edu.iq/index.php/anjs/article/view/2306. [Accessed: 21-Mar-2025].

[16] S. Kumar and A. Gupta, "Enhancement of Random Forest Algorithm Applied in Fake News Detection," World Journal of Advanced Research and Reviews, vol. 18, no. 3, pp. 45-52, 2024. [Online]. Available: https://wjarr.com/content/enhancement-random-forest-algorithm-applied fake-news-detection. [Accessed: 21-Mar-2025].

[17] C. Cortes and V. Vapnik, "Support-vector networks," Machine Learning, vol. 20, no. 3, pp.273–297, 1995.

[18] V. Vapnik, Statistical Learning Theory. Wiley-Interscience, 1998.

[19] B. Schölkopf et al., "New support vector algorithms," Neural Computation, vol. 12, no. 5,pp. 1207–1245, 1999.

[20] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other
Kernel-Based Learning Methods. Cambridge University Press, 2000.

[21] C. W. Hsu, C. C. Chang, and C. J. Lin, "A practical guide to support vector classification,"
Department of Computer Science, National Taiwan University, 2003.

[22] Y. Zhang, H. Zhao, and Y. Chen, "An extension of SVM with an objective function to minimize the influence of outliers," AI & Society, vol. 26, no. 3, pp. 339–348, 2011.

[23] K. Crammer and Y. Singer, "On the algorithmic implementation of multiclass SVMs," Journal of Machine Learning Research, vol. 3, pp. 1057–1080, 2002.

[24] B. Mahesh, "Machine learning algorithms - a review," International Journal of Science and Research (IJSR) ResearchGate Impact Factor, vol. 9, no. 1, 2018, doi: https://doi.org/10.21275/ART20203995.

[25] C. Cortes and V. Vapnik, “Support-vector networks,” Machine Learning, vol. 20, no. 3, pp. 273–297, 1995.

[26] M. T. Owoicho, "Fake News Detection System Using Logistic Regression, Decision Tree and Random Forest," British Journal of Computer Networks and Information Technology, vol. 7, no. 1, pp. 1-10, 2024. [Online]. Available: https://abjournals.org/bjcnit/papers/volume-7/issue1/fake-news-detection-system-using-logistic-regression-decision-tree-and-random-forest/. [Accessed: 21-Mar-2025].

[27] S. Kumar and A. Gupta, "Enhancement of Random Forest Algorithm Applied in Fake News Detection," World Journal of Advanced Research and Reviews, vol. 18, no. 3, pp. 45-52, 2024. [Online]. Available: https://wjarr.com/content/enhancement-random-forest-algorithm-appliedfake-news-detection. [Accessed: 21-Mar-2025].

[28] N. Cristianini and J. Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods. Cambridge University Press, 2000.

[29] V. V. P. Wibowo, Z. Rustam, S. Hartini, Q. S. Setiawan, and J. E. Aurelia, “Comparison between Support Vector Machine and Random Forest for Hepatocellular Carcinoma (HCC) Classification,” 2021 International Conference on Decision Aid Sciences and Application (DASA), pp. 618–622, Nov. 2020, doi: 10.1109/dasa51403.2020.9317083.

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

[31] S. Raphael, Fake News (Random Forest 99% > Naive Bayes 93.8%), Kaggle, [Online]. Available: https://www.kaggle.com/code/rhythmsage/fake-news-random-forest-99-naive-bayes-93-8. [Accessed: Apr. 18, 2025].

[32] E. Çağan Apaydın, Fake News Detector, Kaggle, [Online]. Available: https://www.kaggle.com/code/eminaanapaydn/fake-news-detector. [Accessed: Apr. 18, 2025].

[33] M. Park and S. Chai, "Constructing a User-Centered fake news detection model by using classification algorithms in machine learning techniques," *IEEE Access*, vol. 11, pp. 71517–71527, Jan. 2023. doi: 10.1109/ACCESS.2023.3294613

[34] Bhavik Jikadara, “Fake News Detection,” Kaggle.com, 2023. https://www.kaggle.com/datasets/bhavikjikadara/fake-news detection?resource=download&select=fake.csv (accessed Feb. 25, 2025).

Appendices

- Downloaded Kaggle Zip File.