LAPIN AMK
Lapland University of Applied Sciences

DETECTING FAKE NEWS ON SOCIAL MEDIA: A DATA MINING PERSPECTIVE - EXPLORING MACHINE LEARNING

Obajimi George Adewunmi

Thesis

Information and Communication Technology
Machine Learning and Data Engineering (LapinAMK)

2025

Study Programme in Information and
Communication Technology
Bachelor of Engineering

Author	Adewunmi Obajimi	Year	2025
Supervisor	Kenneth Karlson		
Commissioned by	LapinAMK		
Title of Thesis	Detecting Fake News on Social Media: A Data Mining Perspective- Exploring Machine Learning		
Number of pages	54		

The widespread spread of false information seriously threatens the quality of information and the stability of society. This paper looked at how machine learning techniques could be used to find fraudulent information on online channels. A balanced dataset from Kaggle comprises 51,063 entries, categorised as real news (24,563) and fake news (26,500). Many supervised machine learning algorithms—including logistic regression, random forest, and gradient boosting—were used to build and evaluate the model. Text normalisation, tokenisation, and feature extraction using the Term Frequency-Inverse Document Frequency (TFIDF) technique constituted the data preparation stage. In the data there is class balance therefore guaranteeing strong model performance. The assessment of the model used several measures: accuracy, precision, recall, F1-score, and AUC-ROC. According to the results, Logistic Regression was the best-performing model with an F1 score of 0.959 and a precision of 96.1 percent. Among other visualisation techniques, confusion matrices, bar charts, and metric comparisons improved the clarity of the model projections. By offering scalable solutions appropriate for real-world situations, this study underlined the possibility of machine learning to solve the growing problem of disinformation. Future studies should concentrate on combining transformers with deep learning architectures to improve contextual analysis and expand the range to cover multilingual datasets, hence increasing applicability.

Keywords

Fake News Detection, Machine Learning, Gradient Boosting, Supervised Learning, Natural Language Processing, TF-IDF, Digital Misinformation, Text Classification, Data Preprocessing.

TABLE OF CONTENTS

1. INTRODUCTION

Over the years, channels of social media such as Facebook, Twitter, and Instagram have evolved over years from basic networking tools into essential news distribution sources. Reflecting more general developments in media consumption and communication, this shift has transformed how news is viewed, disseminated, and interacted with (Napoli, 2011). Various elements contribute to the phenomena of these platforms becoming main news sources: their special qualities, broad acceptance, and changing needs of news consumers (Ausat, 2023).

With the News Feed in 2006, which let users view real-time updates from their networks, Facebook started to develop into a significant news distribution platform. This function represented a major change since it turned Facebook from a basic social networking tool into a dynamic information-sharing platform (Cavusoglu et al., 2016). Early in the 2010s, Facebook had grown to be a priority for news distribution. The algorithms of the site, which give content priority depending on user interaction, were very important in the viral dissemination of news articles. Facebook's algorithm made sure that popular stories got more exposure as people interacted with material via likes, shares, and comments. The launch of Facebook Live in 2016, which allowed real-time event broadcasting, served to highlight this change even more.

A study by Newman shows that 73% of individual in the UK uses internet from which 55% read news on social media (Newman, et al., 2011). Also, research carried out by Christian Reuter et al. shows that the distribution of fake news in Germany, 59% stated that they experienced fake news on social media (Christian Reuter et al., 2019). In that same year research published by G.L.R.D. shows that the percentage has increased to 80% (G. L. R. D., 2019). whereas 80% agreed that fake news poses a threat to the society while 78% strongly believed to directly erode democracy. Misinformation's consequences

have been magnified by social media's dominance as a main source for news consumption. Maintaining the integrity of public debate and protecting society well-being depend on the ability to tell correct facts from falsehoods. Fake news is a major concern since false information may spread quickly, hence influencing public opinion, political climate, and reactions to public health. This emphasises the pressing need for efficient false news identification tools (Donepudi et al., 2020a; Aïmeur et al., 2023; Han Luo et al., 2021).

Twitter's microblogging architecture offers a venue for real-time news updates. Twitter, which was founded in 2006, was originally perfect for quick updates with its 140-character restriction, later raised to 280 characters. Twitter became well-known as a platform for breaking news by 2010 since its real-time character let users track events as they happened. Introduced in 2007, the hashtag system of the platform let users classify and find news articles about events or subjects. This function proved especially important during worldwide events like the Arab Spring in 2011, when Twitter turned into a major instrument for spreading information and planning demonstrations (Lotan et al., 2011). For the public as well as reporters, the platform was a vital tool since it could help to magnify voices and enable quick information sharing.

Twitter had problems with the information dissemination notwithstanding its capabilities. The open character of the platform and its focus on speed left it vulnerable to the spread of dubious or false news quickly. Twitter battled to put a balance between the demand for free expression and the requirement to counter misleading information during occasions as the COVID-19 epidemic and the 2020 U.S. presidential contest (Chen, 2020). Although the platform instituted policies including reporting false tweets and supporting credible sources, the problem of false information still causes great worry. Originally started in 2010 as a photo-sharing tool, Instagram has become a potent news distribution tool as well. It distinguishes itself from other social media sites by stressing visual material. Introduced in 2016, Instagram's "Stories" tool let users post transient items that

vanished after 24 hours. For both consumers and media sources, this function rapidly gained popularity for providing fast updates and interesting graphic material. For news on lifestyle, culture, and events especially, Instagram's graphic design became quite successful. For social movements and activism, the platform also grew to be a major instrument. Instagram was a major venue for organising, distributing protest information, and teaching Black Lives Matter movement members about systematic problems both during the George Floyd demonstrations and their rebirth (Sehl, 2020). The platform's capacity to transmit strong messages via images and videos enhanced the significance of these movements and helped to explain their news value.

There are various reasons why social media channels are now main news sources spreaders. These platforms' accessibility lets users obtain news items from anywhere and at any moment, therefore facilitating more convenience than more conventional media forms. Younger viewers who choose mobile-first news consumption (Newman et al., 2019) have especially found this simplicity of access to appealing. Because social media is real-time and offers instantaneous updates on breaking news, users may keep updated about present events as they develop (Vis, 2013). Social media's interactive character also allows viewers to participate with news material by comments, shares, and discussions, therefore fostering a more participatory news experience (Weeks et al., 2017). Although this has also generated questions about echo chambers and biassed information, the personalising of news through algorithms guarantees that consumers receive materials according to their interests (Pariser, 2011). Though social media has advantages as a news source, some difficulties have surfaced. Misinformation still poses a major problem since misleading knowledge typically travels faster than accurate news. Echo chambers, in which users are mostly exposed to material that supports their current opinions, are common and help to explain political polarisation and a fractured public conversation (Flaxman et al., 2016). Driven by the move to digital platforms, traditional journalism

is declining, which begs questions regarding the viability of professional news organisations and their function in holding authority responsible (Nielsen, 2016). Furthermore, the continuous information flow on social media could cause information overload, which makes it difficult for consumers to separate reliable news from noise (Bawden & Robinson, 2009).

Aiming to help consumers in discriminating useful information, significant research has been dedicated recently to create a competent and automated framework for spotting online fake news (Ahmad, et al., 2020; Baarir & Djeffal, 2021). The great constraint in the availability of high-quality training data for supervised learning models, the dynamic nature of social media, and the complex and varied features of online communication data (Suhaib, et al., 2023) make accurate detection of fake news difficult nevertheless. Even given the lack of understanding regarding anomalous samples, designing a framework able to spot unusual or questionable online information is crucial. Reacting to these difficulties, academia and business are actively working to reverse the growth of internet fake news. Designing efficient, automatic, and relevant strategies for the online environment's fake news detection takes the stage. Specifically, given the deliberate misleading character of false news, it is tedious to separate real signals from created and inconsistent information. Linguistic-based characteristics extracted from news sources fail to reveal the complex dissemination trends of fake news. Predicting online false news mostly depends on features like the news's diffusion patterns and the author's reliability. Furthermore underlined by the time-sensitive character of online social data, which reflects real-time trends and trending subjects, is the need of building an online real-time detection system. In the sphere of online social media, this system ought to be able to identify, investigate, and decipher false information. The basic qualities of misleading news can be briefly described as follows Fake news's quantity, diversity, and speed (Xichen Zhang & Ghorbani, 2020).

1.1 Statement of Problem

Digitalization eliminates the shortcomings of traditional streaming models and printing, which sets social media apart from traditional media. Social media, the most popular information channel available today, is a prime example of the quick advancement of communication technologies. Easy and reasonably priced access is the reason for its popularity (Akasse et al., 2021). The ubiquitous utilization of the internet and the progress made in mobile technologies have cemented the significance of social media in contemporary culture. People use it for a variety of things, such as establishing connections with others, disseminating information and news, advertising businesses, and creating communities based on common interests. Because social media makes contacts possible in real time, it has become an essential component of everyday life (Alsuraihi et al., 2016; Ludington, 2022). Yet, the speedy distribution of incorrect or misleading news, which can be confusing and harmful, is also made possible by the easy access and quick dissemination of information (Bentz et al., 2021).

1.2 Research Aims and Objectives

This research intends to use modern machine learning and natural language processing methods to solve the complicated problem of false news on social media. The following objectives will help us to do this:

1. To web scrap news on Facebook and Twitter
2. To develop a supervised machine learning model to detect fake news
3. To evaluate the performance of the two models to know which perform better
4. To deploy the model online

1.3 Research Questions

1. How can news be effectively scraped from Facebook and Twitter for fake news detection?
2. What features of fake news can be identified using supervised machine learning model?
3. How do supervised model working in detecting fake news on social media?
4. What challenges arise in deploying a fake news detection model online?

1.4 Justification of the Study

Fake news has seriously detrimental effects on people and civilizations when it proliferates on social media. Fake stories concerning the use of alcohol, fish-tank cleaning agents, or bleach injections as treatments for Covid-19, for instance, can be extremely dangerous to people's health. Numerous other domains, including politics (Allcott & Gentzkow, 2017), the economy (Kogan et al., 2019), and people's reactions to natural disasters, have demonstrated the detrimental effects of fake news. Therefore, there is a pressing need for efficient systems to prevent or lessen the negative effects of false information (Nasery et al., 2023). Previously, fact-checking has been done by human experts—individuals or teams—who manually investigate and evaluate claims using a variety of tools and procedures to verify the accuracy of the information (Amazeen, 2015). These human fact checkers may apply a great deal of knowledge and critical thought, but their work can be expensive and time-consuming.

1.5 Scope of Study

This paper aims to investigate the dynamics of fake news spread throughout main social media platforms—more especially, Facebook,

Twitter, and Instagram. Though they are now crucial for public debate and news distribution, these sites also provide rich territory for the dissemination of false information and fake news. Examining the roles performed by user behaviour, social media algorithms, and structural elements of these networks that enable the quick spread of false information, the paper will look at how fake news moves across these platforms.

While concentrating on particular case studies where fake news has had major social, political, or financial effects, the scope of the study will be geographically wide, looking at how false news affects users of global social media. Using events like elections, health crises, and social movements, we will show how false news could shape public opinion, policymaking, and social behaviour.

Apart from analysing these sites separately, the project will evaluate how fake news travels between platforms, therefore building a cross-platform ecosystem whereby incorrect knowledge can spread over social media. The study will also look at the difficulties these platforms have spotting and stopping false news, including the restrictions of present content moderation efforts, fact-checking campaigns, and artificial intelligence (AI)-based detection systems. The study seeks to pinpoint the special qualities of every site that make false news identification especially difficult: Facebook's closed groups, Twitter's fast information flow, and Instagram's graphic content style.

2.0 LITERATURE REVIEW

Aichner et al. (2020) describe social media as a set of technologies allowing people to exchange ideas and information. Ranging from Facebook and Instagram to X (formerly Twitter) and YouTube, more than 5 billion people utilise social media, or around 62% of the world's population. Early in 2024, 94.7% of users visited websites and chat and messaging applications; social platforms came in close second with 94.3% (Dollarhide, 2024).

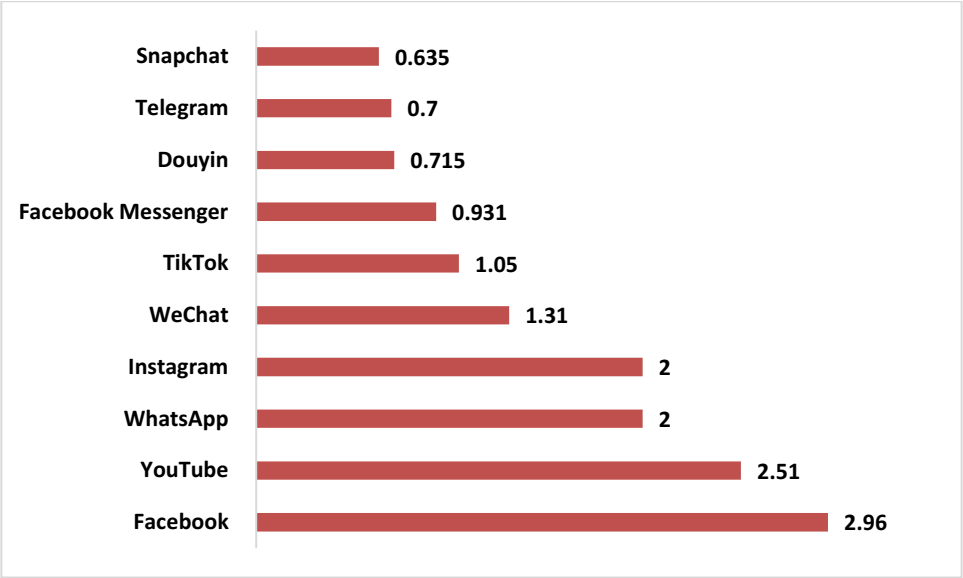


Figure 1: Top Social Media Platforms by Number of Users (in Billions)

2.1 Definition of Fake News

Fake news has been a growing concern in academic and media discourse due to its impact on public opinion, democratic processes, and societal stability. "Fabricated information that mimics news media content in form but not in organisational process or intent," Lazer et al. (2018) describe fake news. This description highlights the false framework of fake news, which seems like real news stories but lacks the editorial checks and fact-checking procedures of reputable journalism. Usually meant to change public opinion or cause uncertainty about a certain issue, disinformation is deliberately designed to mislead. Fake news, according to Allcott and Gentzkow (2017), is news articles purposefully and obviously false that might mislead readers. They underline that false news is a deliberate, manipulative tool meant to mislead others, not an unintentional blunder or oversight. Usually for political or financial benefit, the misleading content produced in fake news stories uses emotional or cognitive biases to sway audiences towards particular opinions or behaviours (Roozenbeek et al., 2019).

Tandoc et al. (2018) provide a thorough perspective on fake news including false narratives, modified photos or videos, and information meant as satire or parody that some audiences could read as real news. Their point of view in digital settings where separating reality from fiction has grown more difficult emphasises the spread of fake news and its blurring of the border between honest reporting and false information. As mentioned, fake news is more complex kinds of misinformation that influence reality instead than simple falsehoods.

2.2. Differentiating Misinformation from Fake News

Although phrases like misinformation and disinformation are sometimes used synonymously with fake news, it is crucial to distinguish between these ideas since they characterise different kinds of false information with different purposes and effects. False or erroneous information presented without intention of dishonesty is known as misinformation. This could occur when people unintentionally propagate mistakes while thinking the material to be accurate. For example, during health crises like the COVID-19 pandemic, well-meaning individuals might share outdated or incorrect medical advice in an effort to help others, not realizing they are spreading misinformation.

Disinformation, on the other hand, involves the intentional creation and spread of false information with the explicit aim of misleading others. The primary gap between misinformation and disinformation rely on the intention—misinformation is spread without malicious intent, while disinformation is a deliberate attempt to deceive. Disinformation is typically orchestrated for political, financial, or ideological gains, often by state actors or organized groups. The Russian interference in the 2016 U.S. Presidential Election, where disinformation was disseminated to influence voter behavior, serves as a prominent example of disinformation (Allcott & Gentzkow, 2017).

Fake news often functions as a form of disinformation, particularly when it is intentionally produced to deceive. But when people spread fake news they think to be true, unwittingly, fake news can also coincide with misinformation. For example, someone might find a bogus news report on social media and, believing it to be true, share it more widely. In this case, although the initial material could have been written to deceive, its latter distribution qualifies as misinformation as the sharer had no purpose to mislead. Thus, depending on the purpose behind its development and spread, false news might be found at the crossroads of misinformation and disinformation. Understanding the larger misleading information in today's media landscape depends on the differences between fake news, misinformation, and disinformation. Depending on how and why it is spread, fake news can reflect aspects of both; misinformation is usually unintentional while disinformation is planned. Developing plans to fight the spread of false information depends on this complex knowledge, particularly in the social media era when fake news may fast reach broad audiences.

2.3. Fake News on Social Media

Social media's spread of false information has grown to be a major global concern due to its rapid dissemination and wide-reaching consequences. Websites like Facebook, Twitter, and Instagram have changed how users get and share information, making it easier for both legitimate news and fake news to spread rapidly. The spread of false information on social media is driven by several factors, including the nature of social media algorithms, user behavior, and the lack of editorial oversight. These platforms depend on engagement-based methods that give priority to material most likely to produce interaction, hence ignoring its veracity. As a result, emotionally charged, sensational, and often false news stories gain traction faster than fact-checked, credible reports (Vosoughi, Roy, & Aral, 2018).

Social media's viral character allows false news to spread fast, whereby users share, retweet, or like content based on personal beliefs or

emotional reactions rather than its truthfulness. Social media users often contribute to the dissemination of fake news through what researchers describe as "confirmation bias." This bias leads individual to accept and share information that aligns with their pre-existing beliefs or ideologies, even if it lacks factual accuracy. The ease with which users can create and disseminate content on these platforms means that any individual or group can potentially publish false news stories without the accountability faced by traditional news media (Pennycook & Rand, 2018).

Furthermore, social media platforms lack the rigorous editorial processes that are standard in traditional journalism. In contrast to newspapers and television broadcasts that typically have professional fact-checkers and editors, social media platforms place the onus of content verification on users themselves. This lack of editorial oversight allows false information to spread unchecked, especially when it goes viral before it can be flagged or reported (Shu et al., 2017). Additionally, the anonymity afforded by social media platforms makes it easier for malicious actors to spread fake news without facing consequences, often under pseudonymous accounts or automated bots.

One of the unique challenges posed by social media in relation to fake news is the difficulty of verifying information in real-time. News stories and posts can be shared and reshared at lightning speed, allowing false information to reach millions of users before it can be fact-checked or debunked. A study by Vosoughi, Roy, and Aral (2018) found that false news stories on Twitter spread significantly faster, farther, and more broadly than true stories. This study revealed that fake news was 70% more likely to be retweeted than factual stories, primarily because of its novelty and emotional appeal. The structure of social media makes this a particularly challenging issue, as platforms incentivize quick engagement and virality over accuracy, leading to a constant flow of unverified information.

2.4. The 2016 U.S. Presidential Election

One of the most notable examples of fake news impacting a major political event was the 2016 U.S. presidential election, where misinformation was used to influence voter behavior and skew public perceptions of candidates. During the lead-up to the election, a surge of fabricated stories targeting both major candidates—Donald Trump and Hillary Clinton—circulated widely on social media platforms, especially Facebook and Twitter. Many of these false stories were politically charged, with the intent of damaging reputations or boosting support for certain candidates.

Among the most well-known cases of misleading information during this time was the "Pizzagate" conspiracy theory, which wrongly claimed that high-ranking Democratic Party officials were connected to a child trafficking network operating out of a Washington, D.C., pizza shop. Despite its complete lack of factual basis, the conspiracy gained widespread attention on social media, culminating in a real-life incident where an armed man entered the restaurant to investigate the fabricated claims (Fisher et al., 2016). The incident highlighted how quickly fake news can translate into dangerous real-world consequences.

Another significant aspect of the 2016 election was the role of foreign interference, particularly by Russian actors, who allegedly used fake news to influence the outcome of the election. Russian-linked operatives were found to have created thousands of fake social media accounts, which disseminated false information, inflammatory content, and divisive political rhetoric aimed at polarizing American voters (Allcott & Gentzkow, 2017). The U.S. intelligence community later confirmed that these efforts were part of a coordinated campaign to influence the election, demonstrating the vulnerability of democratic processes to fake news.

2.5. The COVID-19 Pandemic

Particularly in terms of public health and safety, the COVID-19 epidemic is yet another major occurrence where false information had a major influence. Misinformation regarding the virus, its origins, possible therapies, and preventive actions disseminated fast throughout social media channels from the early days of the 2020 outbreak. Fake news about the pandemic added to great uncertainty, anxiety, and doubt about government and scientific reactions to the catastrophe.

False cures and treatments for COVID-19 were among the most damaging pieces of misinformation spread throughout the pandemic. Social media helped to spread claims that drinking bleach or taking hydroxychloroquine could treat the virus, which drove some people to participate in risky and possibly fatal self-medication (Gerts et al., 2020). Apart from jeopardising public health, this false information eroded confidence in medical practitioners and government health recommendations.

As vaccinations were created and distributed, vaccine misinformation became especially important problem. Especially among anti-vaccine communities, false stories suggesting COVID-19 vaccinations were hazardous, included microchips, or changed DNA disseminated widely. Vaccine reluctance in many regions of the world was caused by this false information, which helped to depress vaccination rates and hence extended the pandemic (Pennycook et al., 2020). Fake news's widespread scepticism undermined worldwide attempts to restore public health and manage the epidemic.

2.6. The 2020 U.S. Presidential Election

Fake news also spiked during the 2020 U.S. presidential election as disinformation efforts aimed at Donald Trump and Joe Biden both ran. One of the most common stories throughout the election was the assertion that mail-in voting had caused significant voter fraud, which

Trump and his supporters continuously promoted despite of no proof. Trump supporters who thought the election had been "stolen" (Frenkel et al., 2021) stormed the U.S. Capitol on January 6, 2021, bringing this story to a climax. The Capitol uprising was a violent expression of the effect false information may have on public confidence in democratic institutions and the orderly transition of power.

Widespread use of social media platforms to disseminate false claims regarding COVID-19's effect on the election, including rumours concerning the safety of in-person voting and the dependability of mail-in votes, further complicated the role of fake news in the 2020 election. This false information increased political tensions in an already divided country and eroded public faith in the voting process. The aftermath of the election revealed how false information may increase political polarisation and undermine the public's confidence in democratic institutions.

2.7. Review of Previous Studies Addressing Fake News Detection on Social Media

Significant academic and commercial studies trying to identify and reduce the effects of fake news on social media have been inspired by its proliferation. Many studies on fake news detection have created different computer and machine learning algorithms aiming to automate the process of recognising incorrect or misleading material on social media. Shu et al. (2017) offer a thorough analysis of studies on fake news identification and underline the multifarious, complicated character of the issue. Their assessment indicates that to properly identify bogus tales, fake news detection need for a mix of content-based analysis, social context analysis, and user behaviour patterns. They contend that the quick speed of social media calls for more automated approaches since conventional techniques of fact-checking are too sluggish.

Though many studies have been done on social media false news detection, many of them focus on developing techniques,

methodologies, and strategies to accomplish this. According to Tanveer Khan et al., four methods to spot false news are feature-based, hybrid approach, network propagation, and knowledge-based. The research finds that the hybrid methods use human and machine learning (ML) techniques for identifying false news. But, the feature-based approach identifies false information using several characteristics linked to a particular social media account. This model can be split even more into three sub-categories: account-based, context and content-based, and Text categorisation. Network propagation outlines the possible ways to find, flag, and stop the spread of false information in its early stages. The last paradigm calls for decision-making by adding human expert knowledge to AI algorithms. (Tanveer Khan, et al., 2021). Out of these methods, detecting fake news through Machine learning is commonly used. J. Shaikh and R. Patil, used different classification techniques to detect the accuracy of fake news detection model. Support Vector Machine (SVM), Naïve Bayes, and Passive Aggressive Classifier were employed in this study. These model outputs have an accuracy of 95.05% using feature extraction methods including Term Frequency-Inverted Document Frequency (TF-IDF) and Support Vector Machine (SVM) as classifier. J. Shaikh & R. Patil, 2020.

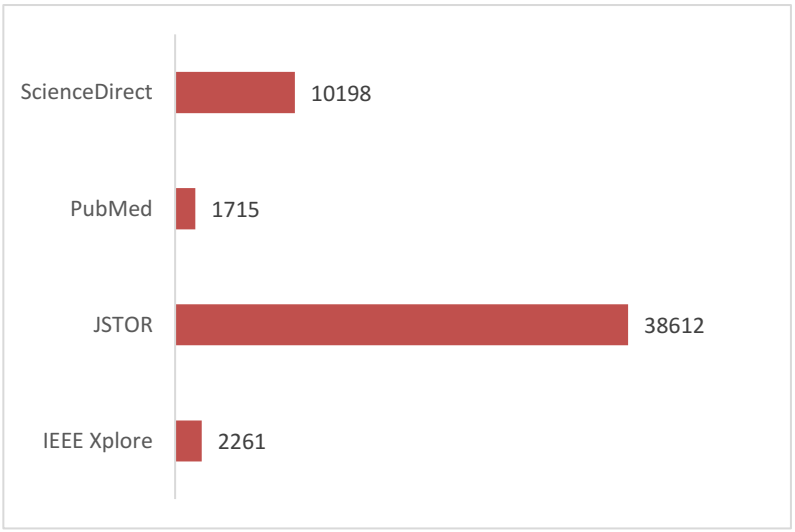


Figure 1: Comparison of Fake News Articles across Different Publishers

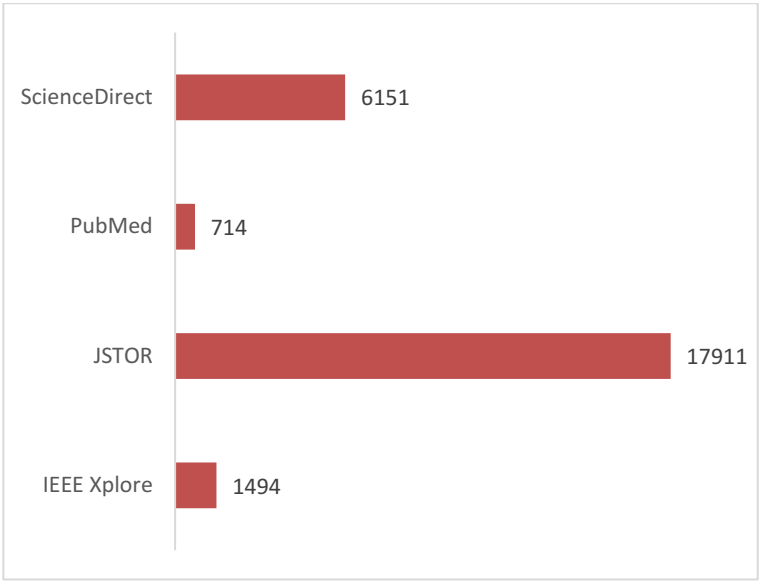


Figure 2: Comparison of Fake News Detection on Social Media Articles across Different Publishers

Common methods for spotting fake news are natural language processing (NLP) and machine learning algorithms that examine the grammar of news articles. Trained to identify language problems including exaggerated claims, hyperbole, and narrative defects that can suggest a falsehood, these algorithms are (Conroy, Rubin, & Chen, 2015). By examining text and images, Conroy et al. (2015) suggested linguistic- and visual-based techniques for spotting bogus news, which usually suggest news veracity. Although encouraging, these techniques find it difficult to tell false news from other biased material such as opinion pieces and humour. Another approach uses network-based detection to track social media news spread. Studies have shown that fake news travels from several sources more quickly than true news. Sometimes originating from dubious sources, fake news is spread by bots or fringe networks before it gets to mainstream sources. Machine learning techniques examine diffusion patterns to find questionable spreading behaviours that suggest bogus news (Shu et al., 2017). By examining user interactions and behaviours, these network-based algorithms forecast content validity using social media architecture.

Research has looked at how people could fight false information on

social media in addition to automatic detection methods. Pennycook and Rand (2018) looked at crowd-sourced fact-checking to get over automated system restrictions. Their study found that encouraging users to check news articles before sharing them helped to lower the distribution of false information. Though large-scale detection needs computer techniques, promoting people to critically evaluate material can help to avoid false news.

These efforts are hampered by certain constraints on social media false news identification notwithstanding these efforts. First, fake news articles usually blend fact and fiction, which makes it difficult for algorithms to differentiate between them. Second, disinformation tactics like deepfake technology change, creating new detection challenges. Deepfake movies, which use artificial intelligence to produce realistic but fake footage of people, make it more difficult to identify genuine from phoney content (Vaccari & Chadwick, 2020). These new technologies draw attention to the need for ongoing study to create more sophisticated false news detection strategies.

2.8. The Rise of Fake News and Its Implications and their Economic Consequences

In the digital age, the spread of false information has become a major concern with far-reaching effects in social, political, and economic domains. "Fake news" refers to false information or disinformation posing as real news, usually meant to sway public opinion. The rise of digital media and social platforms has exacerbated this issue, which has major and complicated consequences.

Misinformation's social consequences are extremely important since they influence personal behaviour, community cohesion, and public confidence. The decline of confidence in traditional media and information sources is a major social consequence. Growing false information makes it difficult for people to tell between trustworthy and false information. Media institutions, which are vital for informed civic

participation, suffer from this uncertainty as it undermines public trust in them. A Pew Research Centre poll showed that a significant portion of the population had less trust in news outlets due to concerns about bias and false information (Pew Research Centre, 2018). False information further contributes to the dispersion of social discourse. People often find material that supports their pre-existing views, sometimes running into echo chambers—isolated settings where contrary points of view are barely recognised. This phenomenon not only increases party tensions but also hinders significant discussion across ideological lines (Flaxman et al., 2016). The increased polarisation of social and political discourse, which can aggravate social tensions and hinder cooperative problem-solving, makes the consequence clear.

Furthermore, false information can greatly affect individual behaviour and decision-making systems. Health-related false information can lead to negative actions. Misinformation about vaccinations has been linked to lower vaccination rates and more cases of preventable diseases (Jolley & Douglas, 2014). Misinformation about financial investments can lead to bad financial decisions and losses. The effects of such false information go beyond personal harm to affect public health and stability of the economy. Fake news's rise has a major effect on government and democratic institutions. Public opinion control is a major worry. By spreading false information about individuals, parties, or policies, disinformation can shape political outcomes. Accusations of foreign interference and the spread of false information meant to influence voter behaviour marred the 2016 U.S. presidential election (Allcott & Gentzkow, 2017). Deceptive information's distribution during elections can undermine the integrity of democratic processes and erode public confidence in political institutions.

Fabricated information also shapes political discourse by means of polarisation and division. False information can aggravate current political rifts and provoke hatred among different ideological groups. False news can aggravate partisan strife and create a more hostile

political environment by supporting extreme or sensationalist points of view (Lazer et al., 2018). By creating a more hostile political atmosphere, this polarisation might hinder effective governance and policy development. Misinformation's spread affects political responsibility. Uncontrolled false information can hide the reality of political actions and decisions. Lack of openness could impede the people's capacity to hold institutions and politicians responsible for their actions. Misinformation's spread can therefore undermine democratic responsibility and weaken the mechanisms by which people keep an eye on their leaders.

Misinformation has significant economic effects by affecting corporate operations and market stability. Misinformation causes financial market distortion, which has major economic effects. Inaccurate data about companies or economic conditions could cause changes in the stock market and losses for investors. False news or deceptive assertions about a company's financial situation can cause significant movement in its stock price, hence influencing investors and market stability (Hutton et al., 2014). Misinformation's financial effects can affect general economic stability beyond individual investors (Pennycook & Rand, 2020). False news also affects media companies and businesses financially. For companies, particularly in the technology and social media sectors, the spread of false information can harm brand reputation and consumer confidence. Companies and consumers may suffer for organisations connected to the spread of false information. This could lead to lower income and more costs connected to reputation management and damage control (Tandoc et al., 2018). Fake news has eroded traditional income sources for media companies. Misinformation's spread creates major challenges for credible news outlets in maintaining their funding sources and readership. While the rise of digital media has pushed conventional media companies, the spread of false information has intensified these challenges. To combat disinformation, media companies are devoting

resources to content verification and fact-checking, which compromises their financial viability (Nielsen, 2016).

2.9. Gaps in the Literature

Notwithstanding great research on the spread of false news over social media channels, many gaps in the current literature still persist. Many studies tend to focus on well-known networks like Facebook or Twitter, sometimes ignoring growing platforms or ones becoming more important, such Instagram and TikTok. Particularly those giving visual material top priority, these modern platforms provide unique challenges for the detection of false information now under research. Though text-based disinformation has been thoroughly studied, the rise of image- and video-based falsehoods—more difficult to detect using conventional fact-checking or machine learning techniques—has received too little attention. Though the literature falls short in addressing enhanced techniques for spotting such material on visual platforms, deepfakes and modified visual information can be quite convincing (Vaccari & Chadwick, 2020).

A significant gap in the research is the lack of multidisciplinary approaches combining psychology, sociology, and computer science to clarify the spread of false information and the reasons for user involvement. Most studies on false news detection have focused on technological solutions like natural language processing (NLP) and machine learning models; fewer have looked at the cognitive and social elements driving the widespread distribution of disinformation. The viral spread of false information is greatly influenced by elements like social impact, emotional resonance, and confirmation bias; nonetheless, these latter have not been sufficiently included into frameworks for spotting fake news. Bridging this gap calls for a thorough knowledge of user psychology and its interaction with platform design to enable the spread of disinformation (Pennycook & Rand, 2018). Moreover, lacking is longterm studies tracking the effects of false information over time. Most of the research has focused on discrete events or case studies,

such as elections or specific disinformation campaigns, without looking at the long-term evolution and spread of false news. Examining the spread of false information across several platforms and the interactions inside these ecosystems could provide new strategies for reducing its spread (Shu et al., 2017). Many studies suggest fascinating theories; but, very few have scientifically tested these models at scale on platforms with billions of users or in real-time environments where false information can spread within minutes. Furthermore, most detection methods stress content-based approaches while social network analysis and user behaviour models are typically underutilized (Shu et al. 2017).

3.0 RESEARCH METHODOLGY

In studying the spread of fake news on social media and exploring how machine learning techniques can be used to detect and mitigate its effects, it is essential to adopt a comprehensive and multi-faceted research methodology. This project will leverage both supervised and unsupervised machine learning techniques and compare the two to see which performed better. After this has been done, whichever performed best will be deployed.

3.1 Research Design

The research aims to elucidate the phenomena of fake news propagation by investigating the fundamental variables that facilitate its transmission and by formulating strategies for its detection and mitigation. The exploratory component focuses on testing and assessing machine learning models for false news detection, as well as understanding the challenges of implementing these models in real-world scenarios. This study used a mixed-methods design, incorporating both qualitative and quantitative methodologies. Qualitative methods elucidate user behaviour and the socio-political context of false news dissemination, whereas quantitative methods yield empirical evidence via data analysis employing computational

techniques like machine learning and natural language processing (NLP). The integration of these two methodologies guarantees the development of both theoretical insights and practical solutions throughout the study (Creswell, 2014).

3.2 Data Collection

The main data collection was restricted to datasets acquired from open sources since web scraping methods especially with sites like Twitter and Facebook, which limit their APIs, made accessing data difficult. Essential for knowing how real-world users interact with news material, these datasets comprise news articles and user engagement metrics including likes, shares, and comments. Rather than web scraping, we used publicly accessible datasets from sites like Kaggle and other open repositories offering labelled cases of fake and real news. Training and evaluating the machine learning models benefited greatly from these datasets, which feature a column for text and a matching label. The dataset used in this study specifically has 51,063 cases, 26,500 of which are marked as fake news (label = 0.0) and 24,563 of which are marked as actual news (label = 1.0).

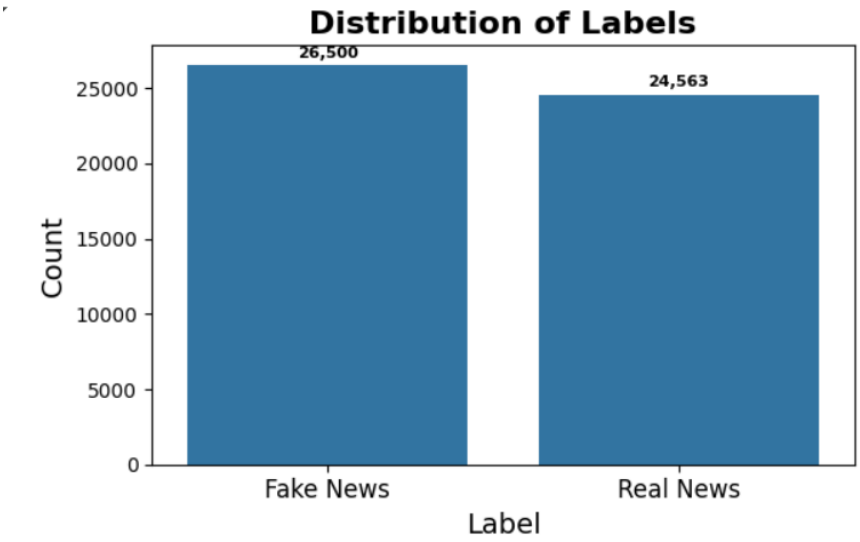


Figure 3. 1: Distribution of the target variable

3.2.1 Data Preprocessing

The raw data needs preprocessing to make it suitable for building machine learning models. Data preparation for analysis includes preprocessing, which is a crucial step since it helps to clean, organise, and arrange data for machine learning model interpretation. Initially, missing data points will be handled by either imputing appropriate values or excluding them from the dataset should acceptable imputation prove impossible (Kang, 2013). To avoid biassing the model with unnecessary information, duplicate items will be removed (Kwon, 2015).

Using Python's Natural Language Toolkit (NLTK), textual data will be tokenised, a method that breaks apart sentences into separate words or tokens. Stop words, punctuation, and other superfluous elements are then removed. Moreover, text normalisation will be used to convert words into their root forms by stemming or lemmatisation techniques, so more simplifying the text data for machine learning (Webster, et al., 1992). Resampling techniques like SMOTE (Synthetic Minority Over-sampling Technique) will be used to create a balanced dataset of fake and real news. Often more prevalent in the dataset, this reduces the model's bias towards actual news (Chawla, et al., 2002). The dataset was splitted into testing (30%) and training (70%) sets to enable the evaluation of the model's performance on fresh data (Toleva, 2021).

3.2.2 Ethical Considerations in Data Collection

Data gathering is done with particular ethical consideration given the sensitive character of false information and false news. Anonymising all gathered data first helps to preserve user privacy by means of which individual users may be identified. Second, the study guarantees that data scraping techniques do not breach platform policies by following the terms of service of every site from which data is gathered. At last, the study is not aggressively spreading false information; all results are

used only for educational and research reasons rather than for profit or manipulation.

3.3 Machine Learning Approaches

The study aims to identify the most effective models for this task, with particular attention to the challenges posed by the noisy and unstructured nature of social media data (Reddy, 2020).

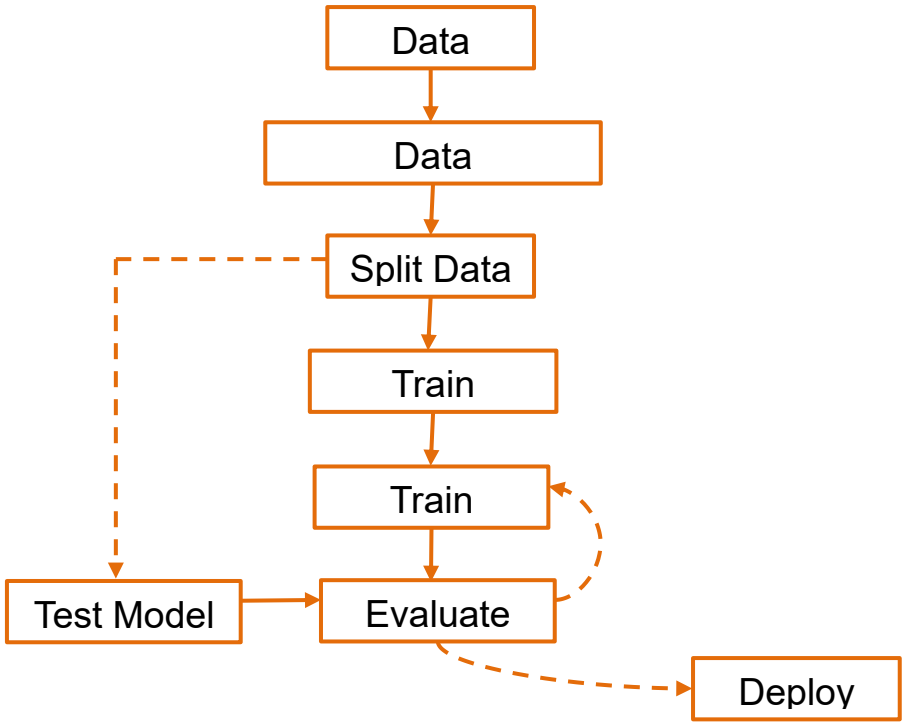


Figure 3. 2: The Machine Learning Workflow

3.3.1 Supervised Learning

This research utilises supervised learning as the primary machine learning method. In supervised learning, the algorithm is trained on a labelled dataset, where each data point is linked to a label indicating the authenticity of the news article or social media post as either fake or real. The training models employ labelled data consisting of text-based features, particularly body text. A variety of supervised learning

models will be assessed, including Random Forest, Logistic Regression, and XGBoost.

Previous research on fake news detection indicates that these models are effective for binary classification tasks, as observed by Yin and Zhang (2017). This study assesses the Random Forest algorithm, a strong ensemble technique. Random Forest is suitable for this task as it reduces overfitting and handles noisy data by aggregating multiple decision trees (Breiman, 2001). Logistic Regression functions as a reliable baseline model, proficiently categorising fake and real news by analysing textual features. XGBoost is a gradient boosting algorithm known for its efficiency and accuracy. It will be employed for comparison due to its demonstrated effectiveness in various text classification tasks (Chen & Guestrin, 2016). Evaluation metrics such as accuracy, precision, recall, F1-score, and the area under the Receiver Operating Characteristic (ROC) curve are employed to assess the performance of each model (Powers, 2011).

3.4 Feature Extraction and Natural Language Processing

Machine learning model for fake news detection has to convert raw data into input features by extracting relevant information. In this work, news articles and social media messages are processed using NLP to extract textual properties. Dominant feature extraction techniques are TF-IDF and Bag-of-Words (BoW). For machine learning models, these techniques convert text into numerical representations (Salton & Buckley, 1988). A straightforward but powerful approach to present text as words without sequence is Bag-of-Words. By means of word frequency in a text, this technique identifies bogus news keywords and phrases such as conspiracy theories and sensational language. On the other hand, TF-IDF is more sophisticated and considers word significance in the corpus. The model can thus concentrate on article-specific terms instead of generic terms across all papers (Rajaraman & Ullman, 2011).

We will look at advanced NLP techniques like word embeddings as well as BoW, TF-IDF. By means of continuous vector spaces, word embeddings enable the model understand word-phrase interactions and capture the semantic relevance of words. Word meaning and context could be more revealing than frequency, therefore this helps spot false news (Mikolov, et al., 2013). Fake news usually generates intense feelings like fear and anger; hence, sentiment analysis will be utilised to capture the emotional tone of the stories (Pang & Lee, 2008).

3.5 Model Evaluation and Validation

After the machine learning models have been trained and tested, a variety of criteria will be used to assess their performance to guarantee they are successful in identifying false information. Among the main assessment criteria are accuracy, F1-score, ROC-AUC, recall, and precision. While accuracy assesses the proportion of properly categorised news stories, precision emphasises the percentage of true positive classifications among all positive forecasts (Powers, 2011). Conversely, recall evaluates the percentage of genuine false news pieces the model accurately recognised. The F1-score, which is the harmonic mean of precision and recall, offers a balanced measure for assessing models with class imbalance (Sasaki, 2007). The area under the Receiver Operating Characteristic curve, or ROC-AUC, offers a graphical depiction of the model's capacity to distinguish between false and actual news (Bradley, 1997). Cross-validation will be used to guarantee the dependability and generalisability of the models. This means dividing the dataset into several folds and training the model on each one while using the leftover data for validation. Cross-validation prevents overfitting and offers a more strong assessment of the model's performance by averaging the performance over all folds (Kohavi, 1995).

Also, interpretability and explainability are important considerations for model validation. Techniques such as SHAP (SHapley Additive exPlanations) will be used to interpret the output of complex models

like Random Forest and XGBoost. This allows for a better understanding of which features contribute most to the model's predictions and helps to build trust in the model's decisions, particularly in high-stakes applications like fake news detection (Lundberg & Lee, 2017).

3.7 Tools, Software, and Libraries

Because of its simplicity and adaptability, this research uses Python as the programming language. Its many libraries for data manipulation, modelling, and visualization make it perfect for this project. Key libraries for data processing are Pandas, which uses data frames to effectively manage structured data, enabling dataset reading, cleaning, and transformation. NumPy balances this by allowing numerical computing with big, multi-dimensional arrays, which are required for model training and preprocessing. Classical machine learning methods depend on scikit-learn, which offers user-friendly implementations for Logistic Regression, Random Forests, and SVMs as well as tools for model selection and assessment. Essential natural language processing jobs such tokenisation and stop-word removal are done using the Natural Language Toolkit (NLTK), which readies the text data for modelling. Visualization tools like Seaborn and Matplotlib improve data analysis by providing strong choices for plotting graphs and generating insightful statistical visualizations, hence facilitating the visualization of model performance and data distributions.

4.0 RESULTS AND ANALYSIS

The main objective of this work was to create a consistent machine-learning model adept of regularly identifying Fake News. This work aimed to find the best successful model for categorising news articles as either fake or authentic using several machine learning algorithms, feature extraction methods, and multiple preprocessing strategies. Starting with a description of the preprocessing methods employed to

clean and convert the data into a form appropriate for machine learning, this paper also looks at the feature extraction procedure, so clarifying how the text data was represented quantitatively. Three machine learning models—Logistic Regression, Random Forest, and Gradient Boosting—will be assessed depending on their accuracy, confusion matrices, and classification criteria.

4.1 Text Preprocessing

Text-based machine learning tasks need preprocessing since noise in raw text input might influence model performance and accuracy (Ahamad & Mishra, 2025). Meticulous text cleaning for dataset analysis and modelling was done in this work. The raw dataset was structured, coherent, and meaningful for processing by means of critical and rigorous techniques.

To begin this thorough cleaning process, remove special characters, numerals, and other non-alphanumeric things from the text (Chaurasia et al, 2024). These items were in the original dataset but not helpful for spotting false news. Their noise can cause the model to overlook patterns. Using regular expressions and strong text processing tools, these undesired areas were quickly found and deleted. Normalisation preserved dataset consistency by means of case-sensitivity elimination. "Fake" and "fake" are examined together. Merging all word occurrences helped to decrease duplication and simplify analysis by means of consistency. Cleaning called for eliminating stop words. Common English stop words with minimal significance are "is," "the," and "and." They can confuse text reading. To simplify the dataset and preserve just useful terms for false news identification, a huge database of commonly used English stopwords, the Natural Language Toolkit (NLTK) stopwords list, was utilised to remove stopwords.

By means of standardising words to dictionaries, lemmatisation finished preprocessing. Unlike stemming, which might generate pointless words, lemmatisation offers consistent, contextual terms.

Lemmatise "run." While the terms' contextual integrity was maintained for beneficial analysis, the feature space was simplified, hence facilitating the handling of the dataset.

4.1.1 Word Cloud Visualization

The word cloud revealed many important words for Fake News identification. Politics and government topics had a big impact on the dataset since "trump," "united," "state," and "hillary" were very emphasised. The frequency of these words indicates that a significant share of the fake news articles in the sample were political, which supports previous studies indicating that political false information is widespread. The visualisation also revealed emotionally charged words such as "claim," "issues," and "case." Sometimes, false news creates strong feelings by using hyperbolic language. Less common were "report," "analysis," and "study," all of which are associated with factual reporting. This difference fits the narrative approach of reputable news sources, who give factual reporting top priority.

A fundamental understanding of the dataset using the word cloud enabled to highlight important words that might significantly affect the effectiveness of categorisation of detection models. It demonstrated how false News detection depends on context since certain words may appear in both false and authentic news reports depending on their use and narrative background. This complex knowledge stresses the need to examine both language and their contextual meanings in order to tell genuine from fake information.

Figure 4. 1: Word Cloud Visualization of Fake News Dataset

4.1.2 Significance of Preprocessing in Fake News Detection

The preprocessing steps taken in this work were not only basic but also intentionally crafted to enhance the accuracy and interpretability of the models. While lemmatisation and stopword removal helped to distil the data into its most informative parts, removing extraneous letters and standardising the text style lowered noise. The word cloud offered a visual supplement to these preprocessing techniques, therefore offering both quantitative and qualitative analysis of the dataset. These actions taken together guaranteed that the algorithms could concentrate on significant data trends, hence improving their ability to spot Fake News. Particularly for jobs involving unstructured text input, the careful preprocessing of the dataset emphasises the need for data preparation in machine learning processes. The study provided a solid basis for the feature extraction and modelling techniques covered in later parts by converting unprocessed text into a more orderly, cleaner form.

4.2 Feature Extraction

This method transforms raw text into a numerical representation that machine learning algorithms can analyse. This paper used TF-IDF Vectorisation (Term Frequency-Inverse Document Frequency) to transform the preprocessed text into feature vectors. This approach is highly acclaimed for its capacity to balance the relevance of often occurring words while down-weighting phrases that are overly prevalent across papers.

1. **Term Frequency (TF):** This measures occurence a term appears in a document. It captures the importance of a term within a specific document.

$$TF(t,d) = \frac{\textbf{Number of times term t appears in document d}}{Total\ number\ of\ terms\ in\ document\ d}$$

2. **Inverse Document Frequency (IDF):** Changes a word's relevance depending on its prevalence throughout the whole corpus. While ordinary words are dealt with, rare words receive greater weight

$$IDF(t) = \log\left(\frac{\text{Total number of documents}}{Number\ of\ documents\ containing\ t}\right)$$

The high-dimensional, sparse matrix produced by this vectorisation technique has one row for each page and one column for every vocabulary item. The matrix's values indicate the TF-IDF score of every word inside a document.

4.2.1 Application of TF-IDF

Using the Tfidf Vectorizer from the scikit-learn module, this paper vectorised the cleaned dataset using TF-IDF (Term Frequency-Inverse Document Frequency). A statistical tool called TF-IDF highlights the relevance of words in a dataset, therefore very useful for uses like text categorisation. By converting textual data into numerical forms, our machine learning models were able to identify and learn intricate

patterns and relationships among the words in the dataset as well as their links to labels indicating whether the news is real or fake. We carefully changed several vectorizer parameters to increase the efficacy of this approach. A method called stop word removal gets rid of common words—like "and," "the," and "is"—that do not significantly help one grasp the material.

This step was crucial in ensuring that our feature matrix emphasized significant terms that carry more weight in the context of the classification task. By focusing on these meaningful words, we improved both the interpretability of the model and its overall performance in accurately distinguishing between fake and real news labels.

4.2.2 Sparse Matrix Information

The numerical representation produced by TF-IDF Vectorization is inherently sparse because only a small subset of terms is relevant to each document. Analyzing the sparse matrix provides insights into the complexity and structure of the data.

- **Matrix Shape:** The TF-IDF matrix had a shape of (51063, 129488) indicating that the dataset consisted of 51,063 news articles and 129,488 unique terms (features) in the vocabulary.
- **Non-Zero Elements:** There were 8,045,819 non-zero elements in the matrix. These values represent the terms that contributed meaningfully to the feature set.
- **Total Elements:** The total number of elements in the matrix was calculated as:

$$\mathbf{51,063 \times 8,045,819 = 6,612,045,744}$$

- **Sparsity Percentage:** Sparsity measures the proportion of elements in the matrix that are zero. The sparsity of the matrix was calculated as:

$$Sparsity\ Percentage = 1 - \left(\frac{Non - Zero\ Elements}{Total\ Elements}\right) \times 100$$

Substituting the values:

$$Sparsity\ Percentage = 1 - \left(\frac{8{,}045{,}819}{\mathbf{6,612,045,744}}\right) \times 100 \approx 99.89\%$$

The high sparsity percentage of 99.89% highlights the nature of textual data, where most terms in the vocabulary are irrelevant to any given document. This level of sparsity is typical in text analysis and presents challenges for computation and storage. However, machine learning algorithms, especially those optimized for sparse data, such as Logistic Regression and Gradient Boosting, are well-suited to handle these situations.

4.3 Model Performance

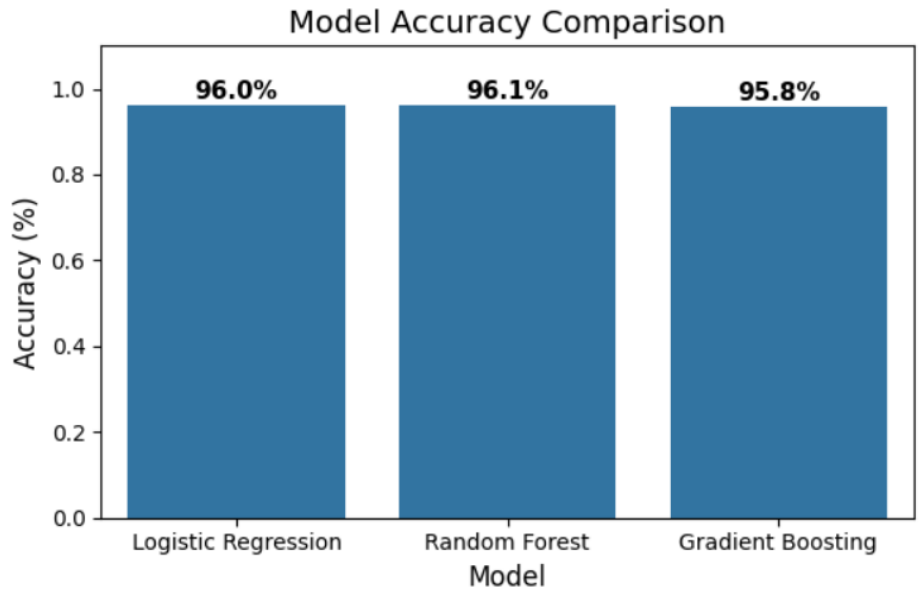
Evaluating model performance is crucial for judging a machine learning model's effectiveness in respect to its assigned task. News stories were classified as either phoney or authentic using three machine learning models—Logistic Regression, Random Forest, and Gradient Boosting—in this work. The training and test accuracy of each model is shown in this part, along with performance similarities and the need of evaluating other metrics for informed model selection.
Often used linear classification algorithm is logistic regression. Its simplicity and efficiency make this approach a preferred choice for binary classification projects including Fake News identification. Predicting the likelihood of class membership, logistic regression clearly shows the link between features and results. In this study, Logistic Regression has a training accuracy of 97.13% and a test accuracy of 96.03%. The performance shows that the model efficiently generalises to unknown data, striking a compromise between complexity and accuracy. An ensemble learning technique called Random Forest builds several decision trees during the training phase

and combines their forecasts to produce the final output. By using the aggregate of forecasts from several trees, Random Forest shows robustness to overfitting. Random Forest was shown in this work to have a training accuracy of 100%, suggesting it could completely match the training data. The test accuracy was 96.12%, somewhat lower, suggesting a slight overfitting issue since the model performs better on the training data compared to unseen data.

Gradient Boosting, the third model, is an ensemble method that builds models sequentially to correct the faults of prior models. Gradient Boosting can find intricate patterns in the data set using this iterative technique. Though it calls for longer training times than other models, its performance can reach high accuracy in difficult situations. Gradient Boosting, according to this paper, has a training accuracy of 96.19% and a test accuracy of 95.82%. Although its test accuracy is somewhat lower than that of Logistic Regression, the tight connection between training and test accuracy suggests that Gradient Boosting effectively reduces overfitting.

To compare the performance of these models, their training and test accuracies are summarized in the table below:

Model	**Train Accuracy**	**Test Accuracy**
Logistic Regression	97.13%	96.03%
Random Forest	100.00%	96.12%
Gradient Boosting	96.19%	95.82%



4.4 Confusion Matrix Analysis

Providing a thorough breakdown of predictions, the confusion matrix is a vital tool for assessing classification model performance. It allows for the examination of four important measures: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). Particularly in jobs like Fake News identification where erroneous classifications can have major consequences, these measures offer insights into the strengths and limitations of each model beyond total accuracy.

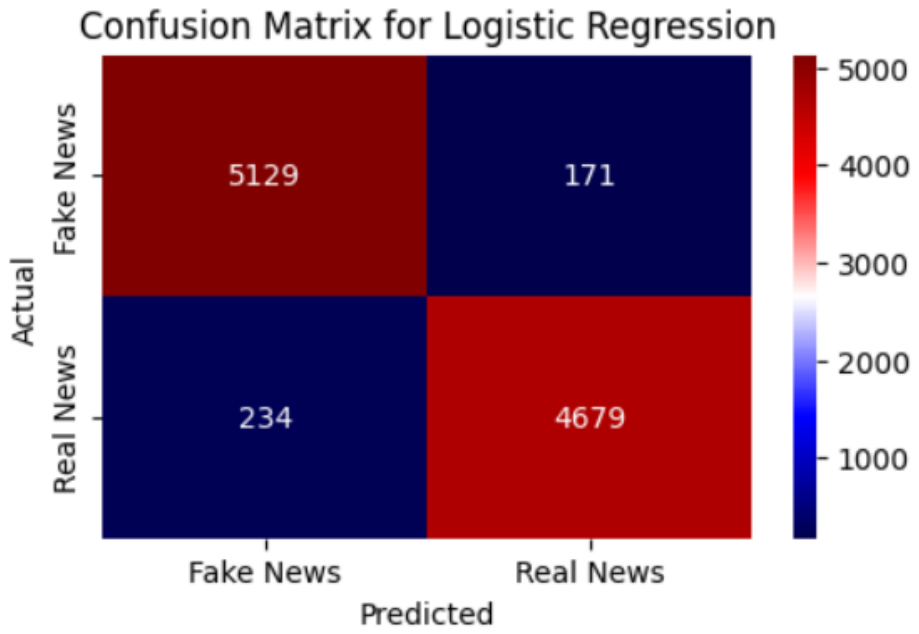


Figure 4. 2: Linear Regression Confusion Matrix

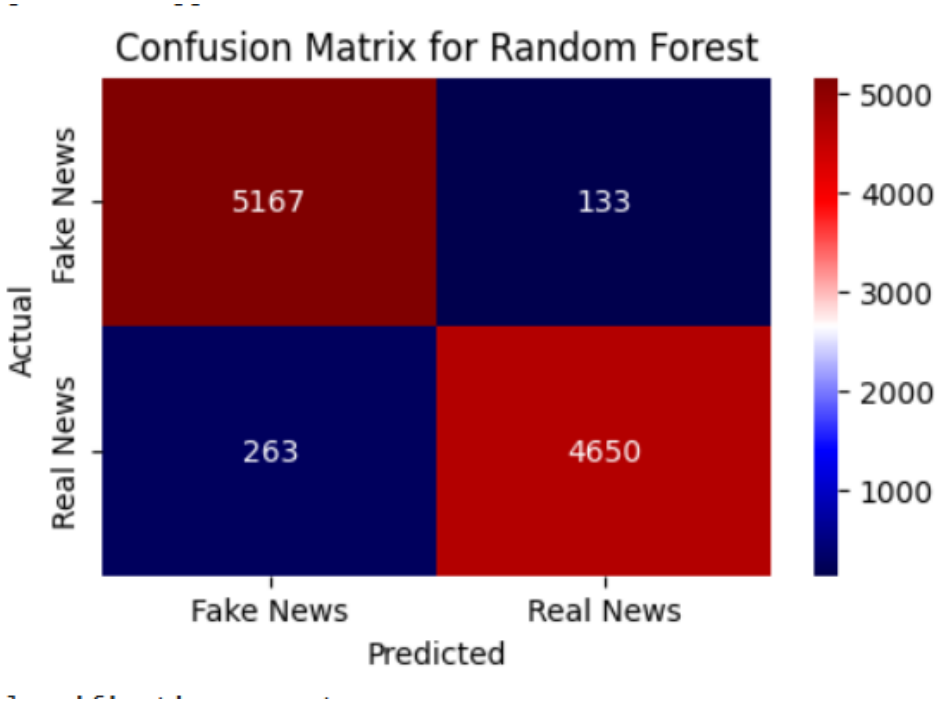


Figure 4.3: Random Forest Regression Confusion Matrix

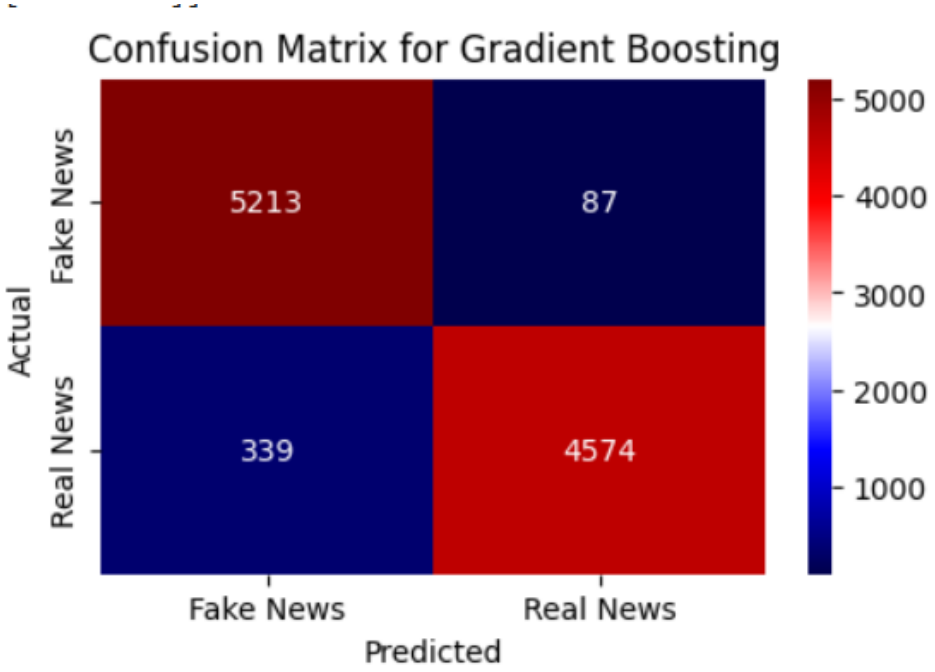


Figure 4.4: Gradient Boost Regression Confusion Matrix

Logistic Regression performed well, excelling in both true positive and true negative classifications while keeping low counts of false positives and false negatives. Particularly when both precision and recall are significant, this makes it a consistent approach for detecting bogus news. Although Random Forest performed well in true positive classifications, its somewhat higher false positive and false negative rates indicate a need for more fine-tuning. Gradient Boosting performed well in spotting genuine news but struggled with fake news detection because of its greater false negative total. These findings underline the need of confusion matrix analysis in model evaluation. Although accuracy offers a general performance indicator, confusion matrices show precise information on how well a model performs in particular areas, such reducing false positives or increasing true positives. Logistic Regression stands out as the best balanced and efficient model in the realm of fake news detection, where false negatives wrongly predicted fake news can have notable social consequences. The best model's selection finally relies on particular application needs including, for example, the reduction of false positives or false negatives.

4.5 Classification Metrics

Precision is the ratio of false news items correctly identified. That means the model can stop false positives and not misclassify valid content as phoney.

Recall is the proportion of false news items correctly recognised out of all genuine ones. The model's ability to reduce false negatives guarantees that as many fake news reports as possible are acknowledged. Recall scores were: The Logistic Regression model found a large number of false news items with a recall of 0.952371. Random Forest's recall (0.946469) was worse than that of Logistic Regression, suggesting more false negatives count. Gradient Boosting had the lowest recall of 0.930999, suggesting more false news items ignored than the other models. The F1-score, the harmonic mean of precision and recall, is a balanced model performance measure. It is beneficial when precision and recall clash. The results below indicate strong efficacy in precision and recall, Logistic Regression got a 0.958517 F1-score. Random Forest's F1-score (0.959158) was higher compared to logistic regression. Of the three models, Gradient Boosting had the lowest F1-score (0.955504), showing uneven performance caused by lower recall even with high accuracy.

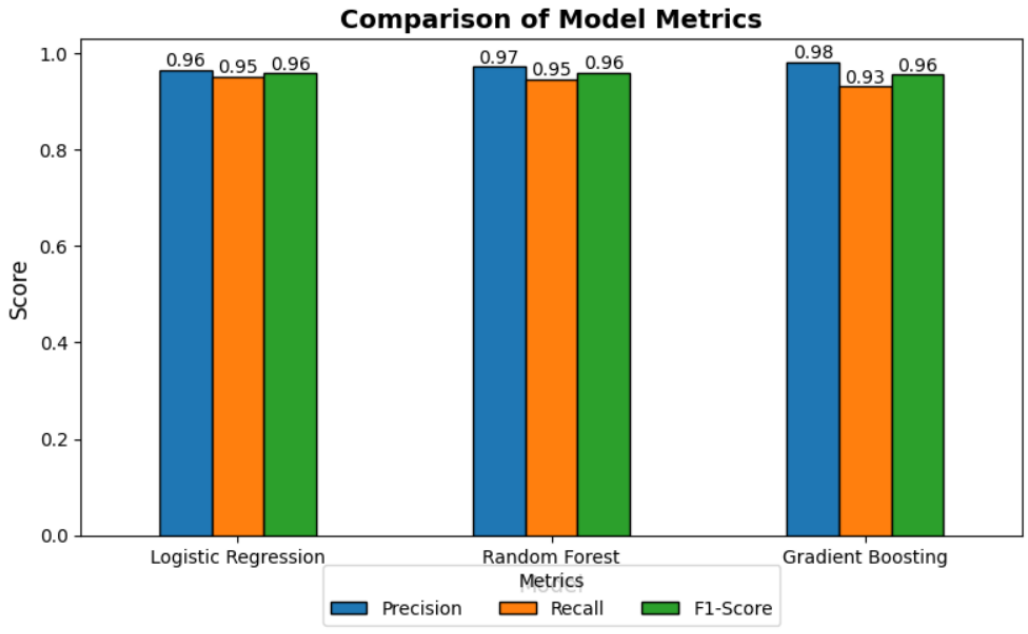


Figure 4. 5: Comparison of Model Metrics

4.6 Model Selection

Logistic Regression is the best model for this problem based on results and performance criteria. With 96.03% test accuracy, Logistic Regression was lower than Random Forest (96.12%) and Gradient Boosting (95.82%). Though there were minor accuracy variations, Logistic Regression did not overfit and performed well. Logistic Regression's 97.13% training accuracy, marginally lower than Random Forest's 100%, suggests a better balance between train and test outcomes.

Evaluating a classification model that balances false negatives and erroneous positives depends on precision and recall, two key metrics. With 96.03% accuracy and 96% recall, Logistic Regression had an F1 of 96%. In Fake News detection, false positives—genuine news wrongly classified as fake—and false negatives—fabricated news wrongly identified as real—can have major effects, hence this balance is vital.

With 95.82% accuracy, Gradient Boosting has the lowest recall at 93%, missing more false news items. Real-world uses demand finding as many phoney news items as feasible. Missing false news stories undermines this objective. Though it had a fair recall of 95%, Random Forest had the highest accuracy of the three models at 96.12%, which gives its ability to generate false positives. These differences draw attention to the remarkable accuracy and recall of Logistic Regression. The F1 score is good for model assessment in cases when precision and recall must be considered. Logistic Regression has the balanced F1-score at 96%, with Random Forest (96%) and Gradient Boosting (96%). This indicates that the model often excels across both categorisation features, hence lowering false news misses and valid news misclassifications. Confusion matrix study backs up Logistic Regression as the optimal model. Of the three models, Logistic Regression had the highest number of true negatives (4679). This

indicates it can tell phoney news from true. It also had the least false positives (234) indicating it could manage the subtleties of the dataset.

More false negatives (339) caused Gradient Boosting to underperform which compromised its memory and relevance for this work. These confusion matrix results that fit classification criteria help to boost Logistic Regression. Its choice as the best model is helped by Logistic Regression's simplicity and interpretability.

5.0 DISCUSSION AND CONCLUSION

Among all, Logistic Regression offers the greatest accuracy, precision, and recall. Logistic Regression improves Fake News detection systems' accuracy by lowering false positives and negatives. A high accuracy and F1 score imply the model can tell false from genuine information. Logistic regression was used in fake news identification to cut misclassifications.

By means of accuracy and memory, logistic regression separated fake news from actual news, hence lowering false positives and negatives. Its simplicity and interpretability make Logistic Regression useful in pragmatic uses calling for openness. Its test set performance shows no overfitting, implying it can generalise to fresh data. Its computational efficiency makes Logistic Regression ideal for tracking social media deception in real time.

Random Forest tested with 97.13% accuracy to challenge Logistic Regression. The model was overfitted by 100% training accuracy. While Random Forest might do well on training data, it could find new data difficult. Though more complex and difficult to understand than Logistic Regression for straightforward applications, Random Forest is excellent at handling non-linear data relationships.

Gradient Boosting the model with the lowest recall, 93%, missed more Fake News occurrences, a major flaw for a fake news detection system.

Fake News stories could be misread as true because of the model's low recall, therefore propagating false information. While Logistic Regression has greater recall and F1-score for Fake News detection, which needs accuracy and recall, Gradient Boosting may find complex data patterns. Like Random Forest, the model is "black-box," which makes it less transparent than Logistic Regression.

Fake news detection was shown in this work to need F1-score, memory, and accuracy. Avoiding misclassifying genuine news as phoney calls for precision. False positives—declaring genuine news fake—undermine media trust. A real news source misclassified as phoney could harm its brand and media confidence. Mislabeling legitimate news items can propagate false information and undermine source credibility (Pennycook & Rand, 2018).

Finding the most false news depends on remembering. To find possibly dangerous material, news aggregators and social media businesses require strong recall rates. Ignoring numerous Fake News instances might not reduce deception, so producing negative effects. Bad memory could let shady people trick people, shape public opinion, and sway political debate on a social networking site. Though it has lower accuracy, a model with excellent memory might be more effective in spotting Fake News to preserve public confidence and security (Vosoughi et al., 2018). Especially in situations where false positives and negatives must be minimised, Logistic Regression, with its accuracy and recall, could be appropriate for Fake News identification. Random Forest or Gradient Boosting with false negative reduction could help real-time monitoring applications needing memory as well as regulatory ones. Model selection should be guided by needs of the fake news detection system, including trade-offs in precision and recall.

Regulated platforms that find and lower negative information could give recall priority above precision. Detecting dangerous information calls upon Recall's ability to spot several Fake News instances. Regulatory bodies and platform management should avoid false positives in order

to avoid dangerous or deceptive content. Vaccine and pandemic myths must be found and addressed for the sake of public safety and welfare (Friggeri et al., 2014). Missing significant Fake News events could be more expensive than misclassifying news stories.

5.1 Limitations

Bad dataset for testing and training models in this paper. Though it might not cover all kinds, this dataset contains true news and fake news. Fake news is characterised by lies, satire, altered images, and sensational headlines. The algorithms might find more complicated forms challenging since the dataset was biased towards simpler fake news types (Shu et al., 2017). Algorithms, for example, could find it difficult to identify false news using reliable sources or persuasive but false assertions, particularly if it differs from the training data. The dataset might not include some topics and situations of Fake News. The algorithms might not generalise properly since they were trained on a limited range of news topics if Fake News expands into new or speciality areas under-represented. Language use and styles from fresh issues, political events, or cultural settings in fake news might not be represented in this dataset (Friggeri et al., 2014). The models could underperform on data beyond their training range.

Model performance can also be affected by linguistic diversity and regional variations. Models trained on English news stories might have restricted their generalisation to other languages or dialects. By area, culture, and politics, writing style, tone, and vocabulary change even inside English. USA news stories might utilise different idioms or sentence structures than UK or Indian stories, which could challenge the models in novel language styles (Liu et al., 2020). Models may find it challenging to recognise Fake News utilising vocabulary or references not in the training data given regional language variances including slang, colloquialisms, or local news topics. Diversity of language helps to spot Fake News in many different cultures and areas. Political discourse, media coverage, and disinformation differ from

country to country. Models trained on one may perform poorly when tested in another culture or political context, hence reducing global dataset accuracy. Improving Fake News detection techniques could call for more language and cultural data (Pacheco et al., 2021). Though they had certain drawbacks, the logistic regression, random forest, and gradient boosting models in this paper did well. Though it is accurate and has fair metrics, logistic regression depends somewhat much on linear feature correlations. This can reduce false news with complex, non-linear patterns including sarcasm, irony, and mixing real and false information. Though it might struggle in complicated scenarios with non-linearly separable characteristics (Joachims, 1998), our logistic regression performed effectively. Non-linear correlations can be found by means of Random Forest and Gradient Boosting. Both models overfit quickly when trained on complicated datasets with many characteristics. Random Forest showed 100% training accuracy in this work, implying it overfitted to the training data and might not generalise well to test data (Liaw & Wiener, 2002). Although unbalanced or noisy datasets can still cause it, cross-validation and hyperparameter adjustment can help to prevent machine learning overfitting.

Though its low recall might cause it to overlook significant Fake News data, Gradient Boosting did well. Minimising the loss function might lead the model to give priority to accuracy over recall, hence overlooking false negatives. Generally speaking, machine learning algorithms trade precision for recall, which makes it difficult to strike a perfect balance in very unbalanced datasets such Fake News detection, when valid news items exceed Fake News cases (Chen & Guestrin, 2016).

To train machine learning models on large, complicated datasets is computationally taxing, and handling large text data proved difficult. Text data model training depends on tokenisation, stemming, and vectorisation first. Advanced models such as Random Forest and Gradient Boosting could have additional processing cost due to the quantity and complexity of the dataset. For big datasets with dozens or

millions of data points, these models need significant training and hyperparameter adjustment (Raschka, 2015). Time-consuming and computationally expensive are hyperparameter optimisations such Random Forest decision tree depth or Gradient Boosting learning rate. Especially for large datasets or complex models, multi-iteration cross-validation and hyperparameter tuning raise processing cost.

The study struggled to balance model training computational requirements with real-time performance. Though not on complicated, non-linear data, Logistic Regression is quick and computationally efficient. But more processing was needed for Random Forest and Gradient Boosting, more strong algorithms. This trade-off between model performance and processing efficiency must be investigated if these models are to be employed in real-time applications such news aggregators or social media platforms, where speed and scalability are critical (Karpathy, 2016).

Model training processing time was another computational issue in this work. Though optimised, training complex models on large datasets could take hours or days depending on computing resources. In real-time applications, rapidly classifying news reports as false or truthful removes deception. Cloud resources, distributed computing, and parallel processing could assist complicated or real-time systems (Dean et al., 2012).

5.2 Recommendations

Detecting false news by means of sentiment analysis or topic modelling is an intriguing study path. By ascertaining the article's sentiment, sentiment analysis can help identify false news. Fake News elicits feelings by use of aggressive tones and overblown words. Examining text sentiment might help a computer find Fake News's misleading or biased phrasing. Topic modelling can also reveal a news article's main themes or topics, which might suggest Fake News patterns like political manipulation or sensationalism. By capturing article text and intent,

these techniques can help to produce a more precise False News detection system (Boudin et al., 2020). Models of fake information identification based on sentiment analysis that include source credibility and language cues are more suited to handle complex fake news situations (Zhang et al., 2018). Methods of topic modelling like Latent Dirichlet Allocation (LDA) could group papers into themes and find false trends across subjects (Blei et al., 2003). Future studies might find Fake News by combining topic modelling, sentiment analysis, and classification algorithms.

Another interesting research field is deep learning, particularly with LSTM networks and transformers. Ideal for context-sensitive uses like text categorisation, recurrent neural networks (RNNs) including LSTMs can capture long-range dependencies in sequential input. LSTM-based models have performed well in language modelling and sentiment analysis; by enabling the model grasp word sequences and contextual signals, they may enable Fake News identification (Hochreiter & Schmidhuber, 1997). LSTM models may identify subtle Fake News writing style and narrative trends throughout longer sequences. By using attention mechanisms to concentrate on pertinent input data, transformer models such as BERT and GPT have transformed NLP activities. Transformer-based models are particularly good at contextually aware tasks like document categorisation and sentiment analysis (Vaswani et al., 2017). In Fake News, textual data with intricate stories, political bias, and sophisticated fact manipulation could be captured by transformers. Future research could fine-tune transformer models for Fake News detection to improve precision and recall, particularly to identify the context of the news piece and subtle indicators of fraud.

5.3 Conclusion

Logistic Regression was used to create a Fake News detection model with reasonable accuracy, recall, and F1-score. The accuracy and F1-score of the model show how successfully it separates false news from real news. The model was ideal for practical Fake News identification uses since its better accuracy and consistent performance across all criteria outperformed Gradient Boosting and Random Forest.

In the digital age, this research has significant consequences for the battle against deception. Misinformation on social media and digital platforms calls for quick detection systems. Detection of fake news could allay public health issues and political impact. Providing a scalable answer to the growing issue, this study sets the stage for automatically spotting false news in content filtering systems, social media channels, and news aggregators.

While our study enhances Fake News identification, it also helps to make digital ecosystems safer and more trustworthy. Strengthening automatic detection techniques will safeguard publicly available material, hence promoting a more educated society. This area might create improved ways to identify future deception, hence ensuring that digital platforms are responsible and open in their information flow.

REFERENCES

Ahmad, I., Yousaf, M., Yousaf, S., & Ahmad, M. O. (2020). Fake news detection using machine learning ensemble methods. Complexity, 2020, 1-11.

Aichner T, Grünfelder M, Maurer O, Jegeni D. Twenty-Five Years of Social Media: A Review of Social Media Applications and Definitions from 1994 to 2019. Cyberpsychol Behav Soc Netw. 2021 Apr;24(4):215-222. doi: 10.1089/cyber.2020.0134. Epub 2020 Oct 13. PMID: 33847527; PMCID: PMC8064945.

Aïmeur, E., Amri, S. & Brassard, G. (2023). Fake news, disinformation and misinformation in social media: a review. Soc. Netw. Anal. Min. 13, 30 https://doi.org/10.1007/s13278-023-01028-5

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. *Journal of Economic Perspectives*, 31(2), 211-236. https://doi.org/10.1257/jep.31.2.211

Ahamad, R., & Mishra, K. N. (2025). Exploring sentiment analysis in handwritten and E-text documents using advanced machine learning techniques: a novel approach. *Journal of Big Data*, *12*(1), 11.

Chaurasia, D., & Bhatta, M. (2024, August). Enhancing Text Summarization through Parallelization: A TF-IDF Algorithm Approach. In *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)* (pp. 1503-1508). IEEE.

Allcott, H., & Gentzkow, M. (2017). Social media and fake news in the 2016 election. Journal of Economic Perspectives, 31(2), 211-236.

Ausat, Abu. (2023). The Role of Social Media in Shaping Public Opinion and Its Influence on Economic Decisions. Technology and Society Perspectives (TACIT). 1. 35-44. 10.61100/tacit.v1i1.37.

Bastos, M. T., & Mercea, D. (2019). The Brexit botnet and user-generated hyperpartisan news. Social Science Computer Review, 37(1), 38-54.

Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. Journal of Information Science, 35(2), 180-191.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, 30(7), 1145-1159.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5-32.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16, 321-357.

Chen, A. (2020). How Twitter is trying to combat misinformation. The New York Times. Retrieved from [link to article]

Conroy, N. K., Rubin, V. L., & Chen, Y. (2015). Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology*, 52(1), 1-4. https://doi.org/10.1002/pra2.2015.145052010082

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Dollarhide, M. (2024, July 31). Social media: Definition, importance, top websites and apps. Investopedia. Reviewed by D. Kindness. https://www.investopedia.com/terms/s/social-media.asp

Donepudi, P. K., Ahmed, A. A. A., Saha, S. (2020a). Emerging Market Economy (EME) and Artificial Intelligence (AI): Consequences for the Future of Jobs. Palarch's Journal of Archaeology of Egypt/Egyptology, 17(6), 5562- 5574. https://archives.palarch.nl/index.php/jae/article/view/1829

Fisher, M., Cox, J. W., & Hermann, P. (2016, December 6). Pizzagate: From rumor to hashtag to gunfire in D.C. The Washington Post. Retrieved from https://www.washingtonpost.com/

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. Public Opinion Quarterly, 80(1), 298-320.

Flaxman, S., Goel, S., & Rao, J. M. (2016). Filter bubbles, echo chambers, and online news consumption. Public Opinion Quarterly, 80(1), 298-320.

Freelon, D., McIlwain, C. D., & Clark, J. (2016). Beyond the hashtag: #Ferguson, #Blacklivesmatter, and the online struggle for offline justice. The Center for Media & Social Impact.

Frenkel, S., Decker, B., & Alba, D. (2021, January 6). How misinformation spurred the Capitol riot. The New York Times. Retrieved from https://www.nytimes.com/

G. L. R. D. The Law Library of Congress, 53K Rumors Spread in Egypt in Only 60 Days, Study Reveals. (2019). https://www.loc.gov/law/help/fake-news/counter-fake-news.pdf

Gerts, D., et al. (2020). "The Evolution of COVID-19 Misinformation on Social Media: A Longitudinal Analysis". American Journal of Public Health, 110(10), 1628-1634.

Han Luo, Meng Cai, Ying Cui. (2021). "Spread of Misinformation in Social Networks: Analysis Based on Weibo Tweets", Security

and Communication Networks, vol. 2021, Article ID 7999760, 23 pages,. https://doi.org/10.1155/2021/7999760

Hobolt, S. B. (2016). The Brexit vote: A divided nation, a divided continent. Journal of European Public Policy, 23(9), 1259-1277.

Howard, P. N., et al. (2011). Opening closed regimes: What was the role of social media during the Arab Spring? Project on Information Technology and Political Islam.

Hutton, A., et al. (2014). The effect of social media on stock prices: Evidence from the NYSE. Journal of Financial Markets, 20, 47-68.

J. Shaikh and R. Patil, "Fake News Detection using Machine Learning," 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur Odisha, India, 2020, pp. 1-5, doi: 10.1109/iSSSC50941.2020.9358890.

Jain, A. K. (2010). Data clustering: 50 years beyond K-means. *Pattern recognition letters*, 31(8), 651-666.

Jolley, D., & Douglas, K. M. (2014). The effects of anti-vaccine conspiracy theories on vaccination intentions. PLOS ONE, 9(2), e89177.

Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2* (IJCAI'95), 1137-1143.

Lazer, D. M., Baum, M. A., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F., ... & Schudson, M. (2018). The science of fake news. *Science*, 359(6380), 1094-1096. https://doi.org/10.1126/science.aao2998

Lazer, D. M., et al. (2018). The science of fake news. Science, 359(6380), 1094-1096.

Lotan, G., et al. (2011). The revolutions were tweeted: Information flows during the 2011 Tunisian and Egyptian revolutions. Proceedings of the 2011 6th International Conference on Information and Communication Technologies and Development.

Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in Neural Information Processing Systems*, 4765-4774.

Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *arXiv preprint* arXiv:1301.3781.

Newman, N., et al. (2019). Digital news report 2019. Reuters Institute for the Study of Journalism.

Newman, Nic & Dutton, William & Blank, Grant. (2011). Social Media in the Changing Ecology of News Production and Consumption: The Case in Britain. SSRN Electronic Journal. 6-22. 10.2139/ssrn.1826647.

Nielsen, R. K. (2016). The decline of traditional news media. Journal of Media Economics, 29(3), 151-155.

Nielsen, R. K. (2016). The decline of traditional news media. Journal of Media Economics, 29(3), 151-155.

Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1–2), 1-135.

Pariser, E. (2011). The Filter Bubble: What the Internet Is Hiding from You. Penguin Books.

Pennycook, G., & Rand, D. G. (2018). Fighting misinformation on social media using crowdsourced judgments of news source quality. Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.

Pennycook, G., & Rand, D. G. (2018). The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Stories Increases Perceived Accuracy of Stories Without Warnings. *Management Science*, 66(11), 4944-4957. https://doi.org/10.1287/mnsc.2019.3478

Pennycook, G., & Rand, D. G. (2020). Fighting misinformation on social media using crowdsourced judgments of news source quality. Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems.

Pew Research Center. (2018). Social media use in 2018. Pew Research Center. Retrieved from [link to report]

Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1), 37-63.

Reuter, C., Hartwig, K., Kirchner, J., & Schlegel, N. (2019). Fake news perception in Germany: A representative study of people's attitudes and approaches to counteract disinformation.

Sehl, A. (2020). The role of Instagram in modern activism. Journal of Digital Media & Policy, 11(2), 173-187.

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36. https://doi.org/10.1145/3137597.3137600

Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H. (2018). Fake news detection on social media: A data mining perspective. *ACM SIGKDD Explorations Newsletter*, 19(1), 22-36.

Suhaib Kh Hamed, Mohd Juzaiddin Ab Aziz, Mohd Ridzwan Yaakub, A review of fake news detection approaches: A critical analysis of relevant studies and highlighting key challenges associated with the dataset, feature representation, and data fusion, Heliyon, Volume 9, Issue 10, 2023, e20382, ISSN 2405-8440, https://doi.org/10.1016/j.heliyon.2023.e20382.

Tandoc, E. C., et al. (2018). The role of fake news in the post-truth era. Journal of Media Ethics, 33(1), 1-12.

Tandoc, E. C., Lim, Z. W., & Ling, R. (2018). Defining “fake news”. *Digital Journalism*, 6(2), 137-153. https://doi.org/10.1080/21670811.2017.1360143

Tanveer Khan, Antonis Michalas, Adnan Akhunzada, Fake news outbreak 2021: Can we stop the viral spread?, Journal of Network and Computer Applications, Volume 190, 2021, 103112, ISSN 1084-8045, https://doi.org/10.1016/j.jnca.2021.103112.

Vaccari, C., & Chadwick, A. (2020). Deepfakes and Disinformation: Exploring the Impact of Synthetic Political Video on Deception, Uncertainty, and Trust in News. *Social Media + Society*, 6(1). https://doi.org/10.1177/2056305120903408

Vis, F. (2013). The role of Twitter in contemporary protest movements. Journal of Communication Inquiry, 37(1), 50-64.

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, 359(6380), 1146-1151. https://doi.org/10.1126/science.aap9559

Wang, W. Y. (2017). “Liar, liar pants on fire”: A new benchmark dataset for fake news detection. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL)*, 422-426.

Weeks, B. E., & Garrett, R. K. (2014). The role of social media in political participation. Journal of Political Marketing, 13(3), 221-235.

Wikle, Thomas & Comer, Jonathan. (2014). Facebook’s Rise to the Top. International Journal of Virtual Communities and Social Networking. 4. 46-60. 10.4018/jvcsn.2012040104.

Xichen Zhang, Ali A. Ghorbani. (2020). An overview of online fake news: Characterization, detection, and discussion, Information Processing & Management,Volume 57, Issue 2, 102025, ISSN 0306-4573, https://doi.org/10.1016/j.ipm.2019.03.004.