

عبدالرحمن غبان

Fake_News_Week_9%2BFinal.docx

 Midocean University

Document Details

Submission ID

trn:oid::2945:406742023

Submission Date

Aug 11, 2026, 7:50 PM GMT+3

Download Date

Aug 14, 2026, 5:35 PM GMT+3

File Name

Fake_News_Week_9+Final.docx

File Size

4.8 MB

32 Pages

6,255 Words

36,144 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (it may misidentify writing that is likely human generated as AI generated and likely AI generated as human generated) so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.



Multimodal Fake News Detection Using Multimodal Large Language Model

Prepared By

Name: Mohammed Salem Alkhulaifi ID No.: 261003052

Name: Abdulrahman Ali Ghibban ID No.: 261001713

Submitted in fulfillment of the requirements for the degree Master of
Science in Artificial Intelligence

At

College of Informatics

Midocean University

UAE

Supervisor Name: Dr. Mohamed Eassa

College of Informatics, Midocean University, Comoros

Email:

Abdulrahman.ghabban@gmail.com

Acknowledgements:

We would like to express our deepest and gratitude to our supervisor, Dr. Mohamed Eassa, whose guidance, encouragement, and unwavering support have been invaluable throughout this research journey. His patience, expertise, and willingness to share his knowledge made every challenge feel like an opportunity to learn. His insight discussions and feedback, particularly on the architecture of Multimodal Large Language Models (MLLMs), played a vital role in shaping this research and the development of the proposed fake news detection framework. I am truly grateful for the time, trust, and confidence he placed in me from the very beginning.

I would also like to sincerely thank my colleagues and friends for being a constant source of support and motivation. Whether through thoughtful discussions, sharing ideas, offering technical assistance, or simply providing encouragement during difficult moments, their contributions have meant more than they know. Their perspectives helped refine the methodology and improve the data processing techniques used in this work, and I am genuinely thankful to have had such supportive people alongside me throughout this journey.

Abstract:

The growth of digital media and social networking platforms has transformed the way people consume and share information. While this transformation has improved access to news and communication, it has also spread misleading and fabricated information. Fake news has become a major challenge because it influences public opinion, damages reputations, and affects decision-making processes in social, political, and economic contexts. Unlike traditional misinformation, modern fake news frequently combines textual and visual content to increase credibility and encourage faster dissemination.

Contents

.....	I
ACKNOWLEDGEMENTS	II
ABSTRACT	III
LIST OF ABBREVIATIONS	IV
LIST OF FIGURES	V
LIST OF TABLES	VI
CHAPTER 1: INTRODUCTION	1
1.1 OVERVIEW	1
1.2 PROBLEM DEFINITION	3
1.3 THESIS OBJECTIVES.....	4
1.4 THESIS CONTRIBUTION.....	5
1.5 THESIS ORGANIZATION	5
CHAPTER 2: BACKGROUND AND RELATED WORK.....	6
2.1 INTRODUCTION.....	6
2.2 BACKGROUND	7
2.3 RELATED WORK	12
2.4 RELATED WORK SUMMARY	18
CHAPTER 3: PROPOSED APPROACH (METHODOLOGY)	21
CHAPTER 4: EXPERIMENTAL RESULTS AND DISCUSSION	22
CHAPTER 5: CONCLUSION AND FUTURE WORK.....	23
REFERENCES.....	24

List of Abbreviations :

Abbreviation Full Term

AI	Artificial Intelligence
LLM	Large Language Model
MLLM	Multimodal Large Language Model
CNN	Convolutional Neural Network
BERT	Bidirectional Encoder Representations from Transformers
NLP	Natural Language Processing
ViT	Vision Transformer
TF-IDF	Term Frequency–Inverse Document Frequency
SVM	Support Vector Machine
RNN	Recurrent Neural Network
LSTM	Long Short-Term Memory
Bi-LSTM	Bidirectional Long Short-Term Memory
GCN	Graph Convolutional Network
MCNN	Multimedia Convolutional Neural Network
AUC-ROC	Area Under the Receiver Operating Characteristic Curve

List of Figures:

Figure	Title
Figure 2.1	Difference Between Machine Learning and Deep Learning
Figure 2.2	CNN Architecture
Figure 2.3	BERT Architecture
Figure 2.4	EfficientNet Architecture

List of Tables :

Table 2.1 Comparison of Previous Fake-News Detection Studies

Chapter 1: Introduction

1.1 Overview

News organizations serve as critical channels for public discourse and global information exchange across political, economic, and environmental domains. However, the authenticity of contemporary news streams is increasingly undermined by intentional falsification. Unchecked misinformation distorts public perception while inflicting measurable reputational and financial harm on individuals, institutions, and state entities. Because modern deceptive content frequently pairs fabricated narratives with manipulated visual assets to increase persuasiveness, establishing automated, multimodal verification systems has become essential for protecting digital information networks

News is one of the most important things that people around the world follow daily. They read the news to know what happens around them in different sectors such as politics, sports, economics, environment, as well as knowing what happens in social life daily [1]. However, this news may contain many problems due to falsification. One of the problems that this news can cause is damage to the reputation of individuals, entities, and countries.

Fake news is news stories made entirely of lies and not based on any real source. Due to the massive spread of newspapers and social media sites filled with news and misinformation, they do not only include textual news but also images and videos [2]. Those who spread misinformation often use fake news to attract viewers and readers by finding a controversial topic or public opinion issue, which makes its spread very rapid, especially if it contains visual or audio content on social media platforms. Therefore, it has become very important to find a way to detect misleading news to protect people's rights, and this is the focus of our research.

Moreover, many people unknowingly read fake news, often during their free time. These sites often do not deal only with text content but also with visual content such as images and short videos, and users, in turn, do not verify the accuracy of the presented content [3]. Perhaps one of the most prominent reasons for the spread of fake news is its frequent transmission among people whose beliefs align with it. In this study, we will create a tool that verifies fake news and real news and then informs the user of the result directly.

Mostly machine learning techniques are often used to determine whether news is fake or real. This technique relies on the headline of the news article; if it contains numbers or symbols that raise suspicion, then detect this is fake news [3]. However, this method has become outdated and has been discovered by fraudsters. Therefore, the headline is written correctly, but the content of the news article is misleading.

In addition, it has become critical to find a tool for detecting text and visual content within news articles to prevent fake news from appearing to users. We will use large language models to process text and images and then categorize the news according to the results produced by the model.

1.2 Problem Definition

Current solutions rely entirely on text-based fake news detection; the current solutions have limited accuracy result, with the development of large language model (LLM), combine the detection of fake news through text and images [1] . This research will detect fake news by multimodal approach for some reasons.

The first reason is the effectiveness of multimodal in dealing with text and visual content. Current models that use deep learning or traditional natural language processing to detect texts lack sufficient accuracy to detect fake news.

The second reason is that fake news creators rely on creating texts and visual content that attract viewers and are difficult to detect using traditional methods. This content is usually used to gain more views or to obtain paid media material. Therefore, fraudsters try to create methods that are difficult to detect, which allows them to earn a lot of money.

Dealing with only one type of input is one of the common problems in models designed to solve the fake news problem. As a result, more than one data format, such as images for misleading news, must be included to make the models effective in solving misinformation problems. The more a model can detect fake news through multiple formats such as text and images, the robust it is.

As a result of this problem, the solution proposed by this thesis lies in fine-tuning existing multimodal large language models (MLLM) to find a solution to fake news. However, existing solutions only deal with news containing textual content, and the solution proposed in this paper will solve the problem of fake news that contains both textual and visual content.

1.3 Thesis Objectives

The main goal of this research is:

- Designing and developing effective solutions for detecting fake news using multimodal large language models (MLLM).

- Making the most of modern technologies to address the problems related to detecting fake news.
- Understanding the effectiveness of these models in distinguishing between fake news and real news.
- Examine the strength of the model by predicting and classifying news into true, misleading content, manipulated content, false connection, imposter content, and satire.
- Designing a tool to reduce fake content on the internet that can handle news containing text and images.

1.4 Thesis Contribution

This research contributes to solving the problems of fake content on the internet by:

- Finding a solution based on multimodal large language models to verify fake news and prevent its widespread dissemination.
- Determining the accuracy of the model by testing it in real time helps to improve the results provided if there are any issues.
- Apply multimodal large language models (MLLM) to determine whether the existing content is real or created by scammers.
- Finding an integrated model that helps reduce fake content by understanding how scammers try to spread misleading content online.

1.5 Thesis Organization

The rest of this thesis is organized as follows.

Chapter 2 discusses the theoretical background and relevant research areas.

Chapter 3 presents the proposed methodology.

Chapter 4 analyzes the experimental results.

Chapter 5 concludes the study and suggests directions for future work.

2. Background and Related Work

2.1 Introduction

The following section will discuss the history and the related work upon fake news detection whether using traditional machine learning techniques or the new large language model. This chapter will highlight previous research using available technologies to detect fake news and examine the impact of such research on reducing the visibility of fake news to users.

There are many studies that have been conducted on detecting fake news, as indicated in chapter 1. This is because this problem is widespread on the internet and social media. Furthermore, the emergence of modern technologies and their ability to handle text-related problems has increased the desire to use them to prevent fake news and the problems associated with it.

However, much of the current work on solving the fake news problem relies entirely on text-based issues. Although text is one of the biggest problems related to misinformation, attention should also be paid to the problems caused by visual content. Therefore, misleading news containing texts, images, or symbols that could harm anyone should be detected using large language models.

In addition to finding solutions for fake news containing texts, this section will focus on studying the research previously conducted in methodology, datasets, and techniques previously used to identify the cons and pons of existing solutions and to anticipate potential problems that may arise in the future.

2.2 Background

Fake news refers to intentionally misleading or fabricated information presented in a format that resembles legitimate news reporting. The growth of social media platforms has significantly accelerated the creation and distribution of misinformation due to fast information sharing and limited content verification mechanisms.

Early fake news detection approaches depended mainly on natural language processing and machine learning techniques. These methods focused on extracting textual characteristics such as word frequencies, writing style, sentiment, and semantic relationships to determine whether news articles were trustworthy.

As multimedia content became increasingly dominant, researchers recognized that textual information alone could not fully represent misinformation patterns. Images are frequently manipulated or intentionally paired with misleading captions to influence user perception.

2.2.1 Deep Learning:

Deep learning utilizes artificial neural networks composed of multiple processing layers to extract feature representations directly from complex input data. Through iterative optimization algorithms, such networks capture structural patterns without requiring manual feature engineering. In misinformation detection, deep neural networks enable unified feature extraction across distinct input formats, allowing text narratives and visual media to be evaluated within a common mathematical space. Deep learning does the exact same thing, just millions of times over until it masters a task. To understand it, it is neural networks that act like digital factories. They take raw data like pixels from an image or raw words from a tweet and process them through layers to find patterns that humans or old statistics might miss and work by assigning a neural network to decode the text, another to decode the image, and a third network to cross-reference them to see if the story actually holds together or if

the text and image are contradicting each other [4]. It is important to first understand the basics of machine learning. Machine learning is a type of artificial intelligence that allows systems to learn and improve from experience, without being explicitly programmed. There are several types of machine learning algorithms, including supervised learning, unsupervised learning, and reinforcement learning. In figure 2.1 below there is the difference between machine learning and deep learning

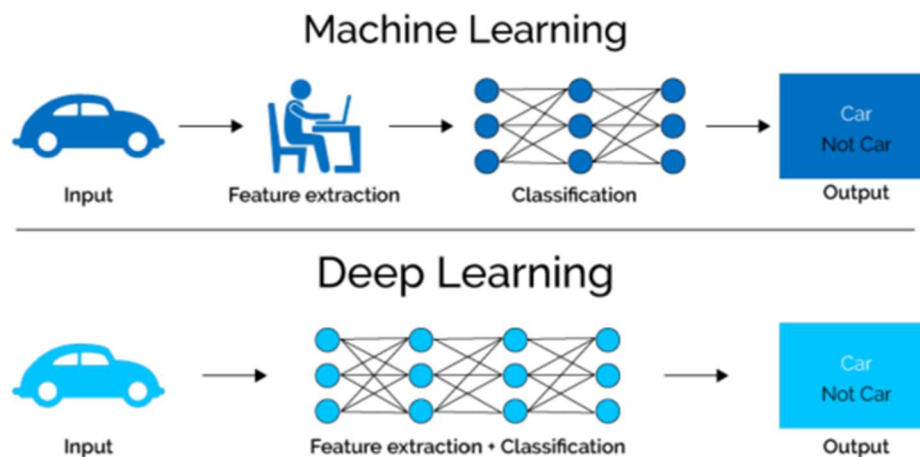


Figure 2.1: Difference between Machine Learning and Deep Learning

Deep learning is a type of supervised learning that uses neural networks with multiple hidden layers to learn representations of the data. These neural networks are composed of interconnected nodes that perform mathematical operations on the data. The layers of the neural network are typically designed to learn increasingly complex features of the data, starting with simple features in the first layer and building up to more complex features in the subsequent layers.

Deep learning introduced more advanced mechanisms capable of learning hidden patterns directly from raw data. Convolutional neural networks became popular for

image analysis, while transformer-based language models improved contextual understanding of text.

A. CNN (Convolutional Neural Network):

Convolutional Neural Networks (CNNs) process spatial data, such as images and video sequences, through parameterized filter operations. Spatial kernels slide across input channels to detect localized visual attributes—ranging from elementary edge gradients in early layers to high-level semantic object structures in deeper layers. For multimodal news evaluation, CNN components inspect images for visual anomalies, digital tampering, or contextual misalignments with paired textual headers.. First, you notice a sharp line, then a curve, then a red patch. Eventually, your brain pieces those small clues together and realizes you're looking at a giant fire truck. A CNN looks at digital images in the exact same way. CNN is widely used in computer vision tasks such as image classification, object detection, and image segmentation. In addition, it excels at capturing hierarchical patterns and spatial relationships in images, making them highly effective in understanding visual information. CNN solves many problems related to computer vision tasks such as facial recognition, medical image analysis, autonomous vehicles, and video analysis. Below it is a figure 2.2 refer to the CNN architecture:

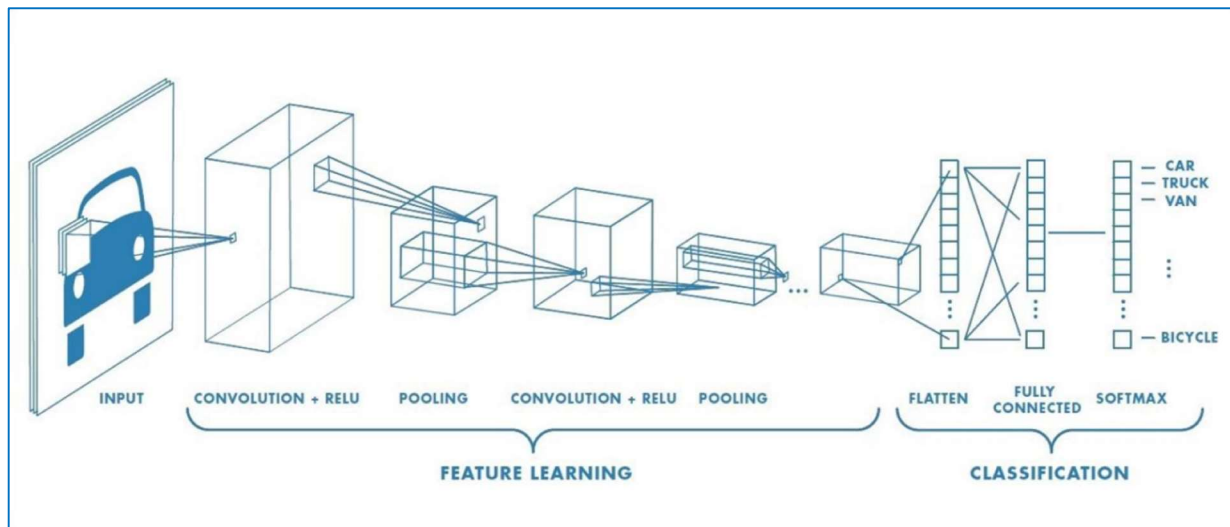


Figure 2.2: CNN architecture

Neurons are the basis of neural networks. The input layer consists of neurons, each of them corresponds to one input feature. Convolution, pooling and RELU all of them are named hidden layers. Hidden layers are known as the layers that exist between the input layer and the output layer. Finally, the output layer is the layer that produces the same number of neurons as the number of classes.

2.2.2 Text approach:

More recently, Large Language Models (LLMs) and Multimodal Large Language Models (MLLMs) have expanded the capability of AI systems by allowing simultaneous understanding of text and visual content [5]. These models provide opportunities to improve fake news detection by analyzing consistency between modalities and generating more informed predictions.

The coolest thing about an LLM is that it's basically a massive accident. They built this system with billions of internal settings, fed it pretty much everything humans have ever written online, and gave it one incredibly simple job: guess the next word in a sentence. Below you will find figure 2.3 which contains BERT architecture, one

of the most efficient models in LLM. BERT is a model developed by Google that understands the context of words in sentences by analyzing words in both directions. BERT often demonstrates high accuracy in text understanding tasks.

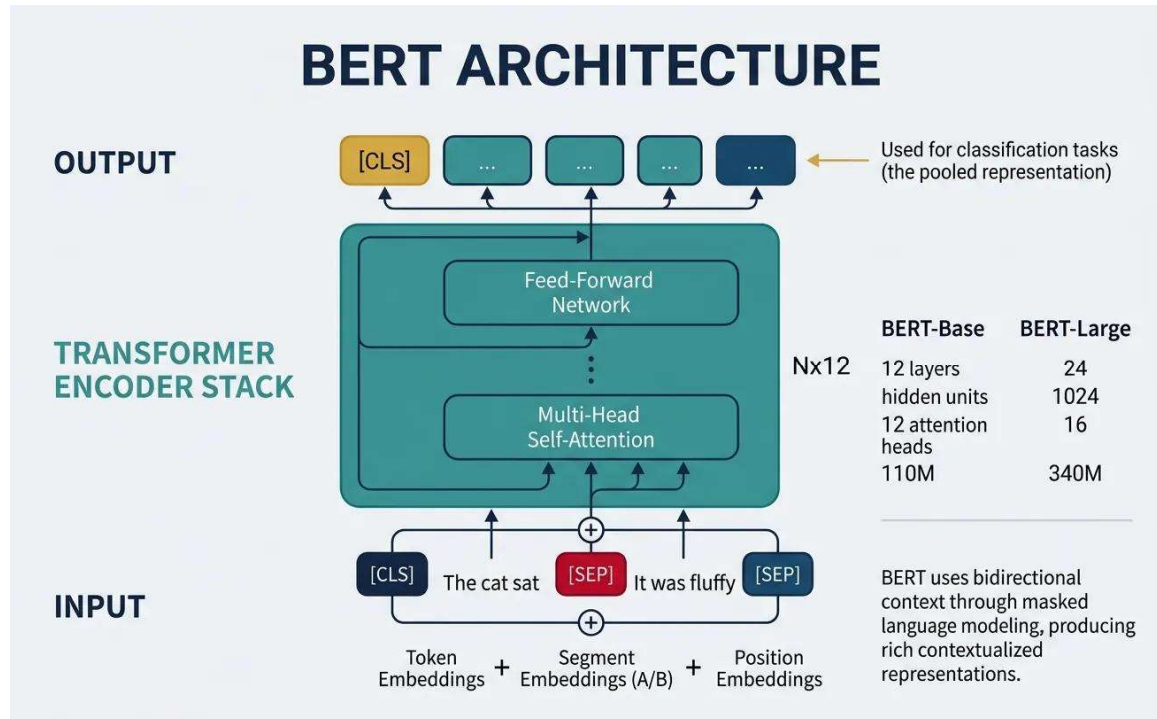


Figure 2.3: BERT architecture

BERT uses a transformer-based encoder to interpret the text then generate contextual representation for each token. Input text named tokens are transformed into embeddings and pushed through the Transformer encoder. A contextual parameter is then generated for each text in the sequence. BERT can read the context from both sides of the word at the same time. Unlike a traditional Transformer that contains both an encoder and a decoder, BERT relies solely on the encoder for language understanding tasks.

2.2.3 Image approach:

One of the CNN families that will be used in this research is EfficientNet. EfficientNet is an architecture that aims to achieve the highest possible accuracy with

the fewest available resources compared to previous models. EfficientNet achieves scalability by incrementally increasing the depth, width, and accuracy of the model based on the compound scaling coefficient ϕ . According to the researchers who developed architecture, this scaling can be achieved by changing each dimension separately. The equation 2.1 below shows the equation of this method:

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \quad (2.1)$$

In addition, find below figure 2.4 EfficientNet architecture will used in this research:

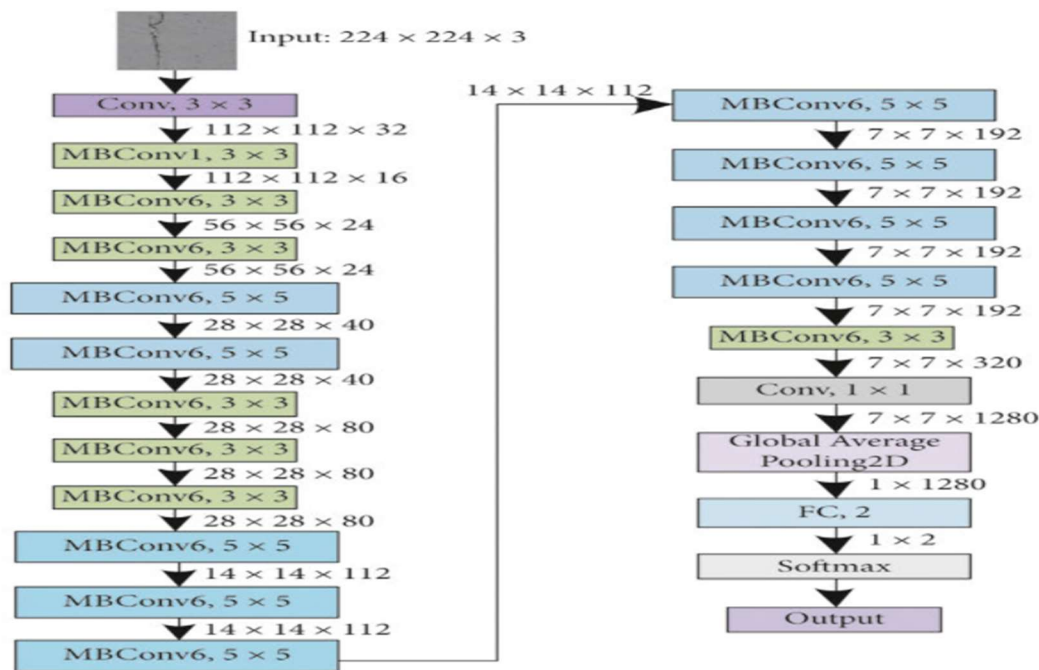


Figure 2.4: EfficientNet architecture

2.3 Related Work:

The reviewed studies demonstrate that fake news detection has evolved from traditional text-based machine learning approaches toward multimodal and transformer-based architectures. Earlier methods achieved reasonable performance but struggled to detect misinformation that combines multiple forms of content.

2.3.1 Text Fake News

[Abraham, Tony Mathew, et al. \(2025\) \[8\]](#) Evaluated classical machine learning classifiers—including Logistic Regression, Naïve Bayes, and Random Forests—on the BuzzFeed and PolitiFact benchmarks, recording a peak accuracy of 94.99%. While their findings illustrate the predictive value of linguistic markers, their experimental design processed textual and visual features through isolated pipelines rather than modeling joint multi-modal interactions.

Addressing cross-modal detection on social network.

[Low, Lawrence Yung Hak, et al. \(2025\) \[14\]](#) proposed a hybrid architecture combining BERT textual representations with visual feature vectors extracted via Xception. Evaluated against the Fakeddit dataset, their model achieved an 86% classification accuracy. Although leveraging pre-trained backbones reduced feature extraction overhead, their evaluation was confined to a single corpus, highlighting the need for cross-dataset validation.

[Embarak, Ossama. \(2025\) \[17\]](#) This research discusses the detection of fake news in Facebook. This research not only deals with detecting fake news through text alone but also analyzes how users interact with this news. The dataset used in this research is Facebook Posts. This study examines user interaction on Facebook through posts and their engagement with them, such as liking, commenting, and sharing. Before training the model, data processing was carried out such as text cleaning, tokenization and lemmatization. The model that worked on combines NLP, Facebook user sentiment, and engagement metrics. The result of the model achieved 89.2% accuracy, 88.4% precision, 86.2% recall, 87.3% F1-score, and an AUC-ROC of 0.91. One of the most important advantages of this research is its ability to detect fake news on Facebook that contains both text and images, rather than just news that contains only text. Also, understanding how users interact with fake news and attempting to analyse it to identify the root of the problem. User interaction may be one of the biggest challenges in this research because it may have been generated by bots.

[Tanaja, Rafael Nicholas, et al. \(2025\) \[18\]](#) In this research, the researchers say that fake news targets younger demographics, therefore it must be detected quickly through machine learning. Researchers claim that fake news spreads faster if there is a global event, citing the coronavirus pandemic as an example. Therefore, the

researchers proposed a rapid detection system to distinguish between fake and real news. The study used the ISOT Fake News Dataset, which was created by University of Victoria. Three different classification methods were used in this research, which are FakeBERT, Style Analysis, Credibility Verification. The results in this study were achieved due to the use of these methods to 90% accuracy, 95% precision for fake news, and 86% precision for true news. The strength of this research lies in its use of FakeBERT, which is based on contextual language understanding, thus providing better performance than machine learning models that rely solely on word frequencies. However, this model is considered very costly from a computational standpoint and slower due to the numerous classifiers.

[Ayyasamy, Ramesh Kumar, et al. \(2025\) \[19\]](#) The research addresses the growing problem of fake news on social media. The researchers explain that although fake news has existed for centuries, social media has greatly facilitated and accelerated the spread of misinformation. Researchers claim that current methods for detecting fake news are inadequate because fake news is created to mimic real news. This study worked on four datasets, which are ISOT, Fakeddit, BuzzFeedNews, and FakeNewsNet. The performance of the developed model yielded different results on the dataset. where the accuracy was 92% on the ISOT dataset, 95% on Fakeddit dataset, 98% on BuzzFeedNews dataset, and 96.5% on FakeNewsNet dataset. This research showed that using deep learning greatly helps in detecting fake news across a diverse dataset.

[M, Visweswaran, et al. \(2024\) \[12\]](#) The authors of this research claim that their main contribution will be in establishing a detection approach for fake news that includes multimedia content. This approach will represent the text as a graph and extract textual features using Graph Convolutional Neural Networks (GCN). In addition, it will extract image features using Vision Transformer (ViT). The dataset that was worked on in this research is the Fakeddit dataset. In this study, the researchers represented the entire text as a graph rather than a sequence of words. They use two deep learning models. The first one is TextGCN which converts text into a graph of documents, words, and weights relationships. The second one is Vision Transformer (ViT) that divides each image into patches. The accuracy of this methodology is 94.17%. The advantage of this approach is its ability to distinguish fake news from real news due to using of multimodal techniques. On the other hand, this approach to dealing with fake news problems is considered costly from a computational standpoint.

[Wang, Jingwei, et al. \(2024\) \[13\]](#) The researcher present that it is not enough to rely solely on detecting fake news that contains only text or images. They developed a multimodal system that analyzes text and images and their consistency to determine whether they constitute fake or real news. The authors worked on three datasets collected from social media platforms they are Weibo, GossipCop and PolitiFact. They combined text features, image features, and LLM reasoning to classify news as fake or real. The accuracy of the model developed in this paper that worked on the three datasets is on Weibo achieved 91.2%, on GossipCop reached 90.5% and on PolitiFact more than 92.5%. One of the most important advantages of this research is its strong performance on multiple datasets. However, there is a problem with implementing this model as it includes a very complex architecture that is difficult to implement in real time.

[Balshetwar, Sarita V., et al. \(2023\) \[11\]](#) The researchers in this study present more on the technical aspect than others, which raised a question for them about whether the quality of data inputs is as important as the choice of artificial intelligence models. In addition, they relied on focusing attention on data cleaning and statistically filling in missing data and then applying machine learning models. Three datasets were worked on ISOT Fake News Dataset, LIAR Dataset, and FakeNewsNet. In this research, the researchers focused on processing missing data and attempting to simulate it before classifying the data and entering it into the models. After processing, all data was transformed using TF-IDF and classified with Naïve Bayes, Passive-Aggressive, and CNN to ensure robust evaluation. The researchers concluded that feature extraction is very helpful in classifying news as fake or real, in addition to multiple imputations, as it helps reduce the error rate in the dataset with missing value. The accuracy of this research reached 90%. The authors also concluded that data processing helps reduce the time spent on computation.

[Gifu, Daniela. \(2023\) \[15\]](#) This paper presents a new perspective on detecting fake news. Researchers claim that fake news is changing people's beliefs and opinions in an unprecedented way. In addition, researchers say that the news recommendation system contributes significantly to the spread of misinformation in an unintentionally. The methodology in this paper is to generate misleading news using LSTM and GPT-2 and then test whether the fake news generated by large language models can overcome fake news detector such as Fakebox. Three datasets were used:

WebText, George McIntire's fake-real-news-dataset, and Covid-19 fake news dataset. Due to the tremendous advancements in LLM technology, the researchers presented a significant improvement between the existing model and the model developed in this research. The models used in this research achieved an accuracy rate exceeding 93%. One of the advantages of this research is the realistic testing of fake news generated by artificial intelligence. One of the main drawbacks of this paper is its reliance on a single tool for detecting fake news, and its limited detection of fake news that contains only text.

[Dharsini, Dr. S. Visnu, et al. \(2023\) \[16\]](#) The research present that the internet gives everyone the opportunity to easily publish information online, which makes fake news spread rapidly on online platforms. In this paper, the researchers presented an approach that combines machine learning such as Naïve Bayes, and Support Vector Machine (SVM) integrated with natural language processing techniques. Because of this approach based on integrating modern technologies, the accuracy of the model presented by the researchers reached 93.5%. One of the most important advantages of this research is achieving very high results by using simple models that are easy to link to existing systems. This research relied on semantic search, which captures the contextual meaning of sentences, instead of relying solely on keywords as in some previous studies. Although the paper is recently published, it lacks the use of some modern techniques such as deep learning, which is considered one of the most effective techniques for detecting fake news. One of the drawbacks of this research is not sharing the dataset publicly.

[Xue, Junxiao, et al. \(2021\) \[7\]](#) Research has shown that, because of the rapid growth of the internet and social media, the spread of fake news has increased significantly. This has caused significant problems and risks to society, politics, and public health. This research relied on four datasets TI-CNN. it contains 20015 news, MCG-FNeWS it is a Chainese dataset, MediaEval the dataset consists of tweets, and each tweet contains either text, an image/video, or social context information, and PolitiFact This data was compiled by experts in the field to verify the real news and distinguish it from fake news. The research used a Multimedia Deep Learning (MCNN) framework that combines BERT for text encoding and ResNet50 for extracting semantic features and image features related to manipulation. This framework reached 95.8% in accuracy. One of the advantages of this research is the use of BERT, which provides a contextual understanding of textual content, thus helping to

understand the semantic features. One of the drawbacks is the use of an older architecture such as BiGRUs instead of Vision transformers or multimodal transformers.

[Roumeliotis, Konstantinos I., et al. \(2021\) \[10\]](#) In this paper, the researchers present to prove which artificial intelligence techniques are most effective for detecting fake news, considering two important factors: model performance and computational cost. The WELFake Dataset from Kaggle was used in this research, which consists of over 72,000 articles categorized as either real or fake news. In this study, researchers tested several techniques, including CNN, BERT, GPT-4o, and GPT-4o-mini, to determine the best methods for balancing accurate fake news detection with minimizing the computational cost of fake news detection. Among the findings discovered in this study, GPT-4o and GPT-4o-mini achieved 98.6% accuracy after fine-tuning, outperforming BERT (97.5%) and largely exceeding CNN performance (58.6%). The researchers in this study concluded that fine-tuning is crucial for achieving optimal performance using GPT models to detect fake news.

[Berger, Lara Marie, et al \(2025\) \[20\]](#) This paper wants to identify which factors are most effective in reducing the spread of fake news on social media. The study found that education about the dangers of fake news was more effective than factchecking because it helped participants distinguish between fake and real information, not just posts that had been explicitly corrected. The dataset used in this study is two-wave survey dataset. The researchers in this paper utilized a randomized controlled trial with two waves of surveys to compare a fact-checking intervention, a media literacy intervention, and a control group. They measured changes in beliefs, knowledge, and attitudes, and then analyzed the results using ordinary least squares of regression, Intention-to-Treat analysis Technique, and additional robustness tests. However, the accuracy of this approach isn't explicit. This efficient design gave them knowledge not only of whether the interventions were successful, but also how long their effects lasted and for whom they were most effective.

[Alshuwaier FA, Alsulaiman FA \(2025\) \[21\]](#) This paper explains that the rapid growth of digital media has made it easy for people to access and share news without verifying its credibility. The dataset used in this study is the Kaggle "Fake News Challenge. This study conducted several experiments on machine learning algorithms such as Naïve Bayes, Support Vector Machine (SVM), and Random

Forest. In this study, scientists also attempted to use deep learning algorithms such as convolutional neural networks CNN, recurrent neural networks RNN, and bidirectional long-term memory networks (BI-LSTMs). The accuracy achieved here is more than 98%. This study demonstrated that the effectiveness of each model depends on four key factors dataset quality, Feature engineering, Training process, and Evaluation metrics.

[Khanam, Z., et al. \(2021\) \[9\]](#) presented to create machine learning-based tools that can help detect fake news. This study aims to create a tool that integrates with other systems such as news websites, web browsers, and social media platforms to classify news as fake or real. This research used a LIAR-PLUS dataset related to political information whose source was verified. Each statement in the dataset was classified as True or Fake. For this reason, supervised learning technology was used to classify data. Instead of using a single algorithm, the researchers relied on comparing more than one algorithm. For example, Naive Bayes, Random Forest, and Decision Tree after using all these algorithms, they selected the best algorithm in terms of performance. The researchers state that the accuracy achieved more than 75%. One of the advantages of this research is the use of several machine learning algorithms instead of relying on just one. One of the drawbacks of this research is its reliance solely on sentences to detect fake news, without considering other factors such as the news publisher or the images included in the news.

2.4 Related Work Summary

Recent multimodal approaches show promising improvements because they analyze interactions between text and images simultaneously. However, many existing solutions focus mainly on classification performance and provide limited investigation into model robustness and practical usability.

Based on these observations, this research proposes a multimodal large language model framework designed to improve fake news detection by combining textual and visual analysis into one integrated system.

Ref	Author	Year	Technique Used	Accuracy	Dataset	Comparison with This Research
7	Xue, Junxiao, et al.	2021	Multimedia Deep Learning (MCNN) with BERT + ResNet50 + BiGRU	95.8%	TI-CNN, MCG-FNeWS, MediaEval, PolitiFact	Evaluates multimodal consistency using MCNN; our model extends this via fine-tuned MLLM semantic reasoning
8	Abraham, Tony Mathew, et al.	2025	Logistic Regression, Naïve Bayes, Random Forest, NLP and data analysis techniques	94.99%	BuzzFeed and PolitiFact	Analyzes classical ML text classification; our framework integrates end-to-end vision-language backbones.

9	Khanam, Z., et al.	2021	Naïve Bayes, Random Forest, Decision Tree (supervised learning)	More than 75%	LIAR-PLUS	Finding a way to prevent fake news from appearing on social media
10	Roumeliotis, Konstantinos I., et al.	2021	CNN, BERT, GPT-4o, GPT-4o-mini (fine-tuned)	98.6% (GPT-4o / GPT-4o-mini), 97.5% (BERT), 58.6% (CNN)	WELFake Dataset (Kaggle)	Focuses on single-sentence text statements; our architecture incorporates image context to prevent single-modality bypasses
11	Balshetwar, Sarita V., et al.	2023	TF-IDF + Naïve Bayes, Passive-Aggressive, CNN, multiple imputation for missing data	90%	ISOT Fake News Dataset, LIAR Dataset, FakeNews Net	Handling and thoroughly testing multiple datasets once

12	M, Visweswaran, et al.	2024	TextGCN + Vision Transformer (ViT) multimodal framework	94.17%	Fakeddit	Studying the requirements for building multimodal to detect fake news
13	Wang, Jingwei, et al.	2024	Multimodal (Text Features + Image Features + LLM Reasoning)	91.2% (Weibo), 90.5% (Gossip Cop), >92.5% (PolitiFact)	Weibo, GossipCop, PolitiFact	Knowing how to work on multiple datasets simultaneously
14	Low, Lawrence Yung Hak, et al.	2025	BERT + Xception + Dense Layer	86%	Fakeddit	Knowing how to handle the data we selected in the thesis
15	Gifu, Daniela. .	2023	LSTM + GPT-2 + Fakebox	93%	WebText, Fake-Real-News Dataset, COVID-19 Fake News Dataset	Understanding how to make this project viable in real time
16	Dharsini, Dr. S. Visnu, et al.	2023	Naïve Bayes + SVM + NLP + Semantic Search	93.5%	Not shared publicly	Deep learning benchmark for fake

						news classific ation
17	Embarak, Ossama.	2025	NLP + User Sentiment + Engagement Metrics	89.2%	Facebook Posts	Understa nding the basics of how detection works on social media
18	Tanaja, Rafael Nicholas, et al.	2025	FakeBERT + Style Analysis + Credibility Verification	90%	ISOT Fake News Dataset	New technolo gies used to solve the problem of fake news
19	Ayyasamy, Ramesh Kumar, et al.	2025	CNN	92% (ISOT), 95% (Fakeddi t), 98% (BuzzFe edNews) , 96.5% (FakeNe wsNet)	ISOT, Fakeddit, BuzzFeedN ews, FakeNews Net	Baseline CNN architect ure for fake news detection
20	Berger, Lara Marie, et al.	2025	OLS regression, Intention-to- Treat analysis	Not sharing publicly	Two-wave survey dataset	Understa nding the problem to be solved
21	Alshuwaier FA, Alsulaiman FA.	2025	Naïve Bayes, SVM, Random	Not explicitl	Not specified	A compreh ensive

			Forest, CNN, RNN, BI- LSTM	y reported		study to identify the techniqu es that help solve the problem
--	--	--	----------------------------------	---------------	--	--

References

- [1]. Tanwar, Vidhu, and Kapil Sharma. "A Review on Enhanced Techniques for Multimodal Fake News Detection." *Recent Innovations in Computing*, edited by Pradeep Kumar Singh et al., vol. 701, Springer Singapore, 2021, pp. 767–77. DOI.org (Crossref), https://doi.org/10.1007/978-981-15-8297-4_61.
- [2]. Giachanou, Anastasia, et al. "Multimodal Fake News Detection with Textual, Visual and Semantic Information." *Text, Speech, and Dialogue*, edited by Petr Sojka et al., vol. 12284, Springer International Publishing, 2020, pp. 30–38. DOI.org (Crossref), https://doi.org/10.1007/978-3-030-58323-1_3.
- [3]. Choudhury, Deepjyoti, and Tapodhir Acharjee. "A Novel Approach to Fake News Detection in Social Networks Using Genetic Algorithm Applying Machine Learning Classifiers." *Multimedia Tools and Applications*, vol. 82, no. 6, Mar. 2023, pp. 9029–45. DOI.org (Crossref), <https://doi.org/10.1007/s11042-022-12788-1>.
- [4]. For the Comprehensive Paradigm Shift (Deep Learning → LLMs)
Zhou, J., & Wang, H. (2026). The Paradigm Shift: A Comprehensive Survey on Large Vision Language Models for Multimodal Fake News Detection. arXiv preprint arXiv:2601.15316.
- [5]. For Integrating LLM Explanations with Deep Learning
Wang, J., Zhu, Z., & Liu, C. (2024). LLM-Enhanced multimodal detection of fake news. *PLOS ONE*, 19(10), e0311241. <https://doi.org/10.1371/journal.pone.0311241>
- [7] Xue, Junxiao, et al. "Detecting Fake News by Exploring the Consistency of Multimodal Data." *Information Processing & Management*, vol. 58, no. 5, Sept. 2021, p. 102610. DOI.org (Crossref),<https://doi.org/10.1016/j.ipm.2021.102610/>
- [8] Abraham, Tony Mathew, et al. "Leveraging Data Analytics for Detection and Impact Evaluation of Fake News and Deepfakes in Social Networks." *Humanities and Social Sciences Communications*, vol. 12, no. 1, July 2025, p. 1040. DOI.org (Crossref),<https://doi.org/10.1057/s41599-025-05389-4>.
- [9] Khanam, Z., et al. "Fake News Detection Using Machine Learning Approaches." *IOP Conference Series: Materials Science and Engineering*, vol. 1099, no. 1, Mar. 2021, p. 012040. DOI.org (Crossref), <https://doi.org/10.1088/1757-899X/1099/1/012040>.
- [10] Roumeliotis, Konstantinos I., et al. "Fake News Detection and Classification: A Comparative Study of Convolutional Neural Networks, Large Language Models, and Natural Language Processing Models." *Future Internet*, vol. 17, no. 1, Jan. 2025, p. 28. DOI.org (Crossref), <https://doi.org/10.3390/fi17010028>.

- [11] Balshetwar, Sarita V., et al. "Fake News Detection in Social Media Based on Sentiment Analysis Using Classifier Techniques." *Multimedia Tools and Applications*, vol. 82, no. 23, Sept. 2023, pp. 35781–811. DOI.org (Crossref), <https://doi.org/10.1007/s11042-023-14883-3>.
- [12] M, Visweswaran, et al. "Synergistic Detection of Multimodal Fake News Leveraging TextGCN and Vision Transformer." *Procedia Computer Science*, vol. 235, 2024, pp. 142–51. DOI.org (Crossref), <https://doi.org/10.1016/j.procs.2024.04.017>.
- [13] Wang, Jingwei, et al. "LLM-Enhanced Multimodal Detection of Fake News." *PLOS ONE*, edited by Hikmat Ullah Khan, vol. 19, no. 10, Oct. 2024, p. e0312240. DOI.org (Crossref), <https://doi.org/10.1371/journal.pone.0312240>.
- [14] Low, Lawrence Yung Hak, et al. "Multimodal Fake News Detection Combining Social Network Features with Images and Text." *Proceedings of the 37th Conference on Computational Linguistics and Speech Processing (ROCLING 2025)*, edited by Kai-Wei Chang et al., Association for Computational Linguistics, 2025, pp. 266–76. ACLWeb, <https://aclanthology.org/2025.rocling-main.28/>.
- [15] Gifu, Daniela. "An Intelligent System for Detecting Fake News." *Procedia Computer Science*, vol. 221, 2023, pp. 1058–65. DOI.org (Crossref), <https://doi.org/10.1016/j.procs.2023.08.088>.
- [16] Dharsini, Dr. S. Visnu, et al. "AN AI SYSTEM FOR FAKE NEWS DETECTION ON SOCIAL MEDIA." *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, vol. 07, no. 11, Nov. 2023, pp. 1–11. DOI.org (Crossref), <https://doi.org/10.55041/IJSREM26757>.
- [17] Embarak, Ossama. "Deep Learning for Fake News Detection: Analysing Facebook's Misinformation Networks." *Procedia Computer Science*, vol. 265, 2025, pp. 199–208. DOI.org (Crossref), <https://doi.org/10.1016/j.procs.2025.07.173>.
- [18] Tanaja, Rafael Nicholas, et al. "Fake News Detection Using Machine Learning: Integrating FakeBERT Classification, Style Analysis, and Credibility Verification." *Procedia Computer Science*, vol. 269, 2025, pp. 1067–76. DOI.org (Crossref), <https://doi.org/10.1016/j.procs.2025.09.048>.
- [19] Ayyasamy, Ramesh Kumar, et al. "A Hybrid Deep Learning Framework for Fake News Detection Using LSTM-CGPNN and Metaheuristic Optimization." *Scientific Reports*, vol. 15, no. 1, Nov. 2025, p. 41522. DOI.org (Crossref), <https://doi.org/10.1038/s41598-025-25311-x>.
- [20] Berger, Lara Marie, et al. "Debunking 'Fake News' on Social Media: Immediate and Short-Term Effects of Fact-Checking and Media Literacy Interventions." *Journal of Public Economics*,

vol. 245, May 2025, p. 105345. DOI.org
(Crossref),<https://doi.org/10.1016/j.jpubeo.2025.105345>.

[21] Alshuwaier FA, Alsulaiman FA. Fake News Detection Using Machine Learning and Deep Learning Algorithms: A Comprehensive Review and Future Perspectives. Computers. 2025; 14(9):394.
<https://doi.org/10.3390/computers14090394>