Fake News Detection Using Large Language Models

Lynda Abbas, Zara Afreen, Elena Petukhova, Nafiseh Vahidian

I. Introduction

In today's digital era, the widespread use of the internet and social media has dramatically transformed the way information is created, disseminated, and consumed. While these advancements offer unprecedented access to knowledge and communication, they have also facilitated the rapid spread of misinformation, particularly fake news. Fake news can originate from disinformation (intentionally fabricated content for specific purposes) or evolve as misinformation (false content shared unknowingly by individuals who believe it to be true). Psychological factors, such as confirmation bias, further exacerbate the issue; people tend to accept and share information that aligns with their personal beliefs, regardless of its factual accuracy. Behavioral studies suggest that exposure to fake news increases the likelihood of belief in the content, and even when individuals recognize the content as false, they may still share it to affirm social identity, signal political alignment, or elicit emotional reactions from their network [3] .

An MIT study that tracked 126,000 stories revealed an alarming trend, fake news stories are approximately 70% more likely to be retweeted on Twitter than factual stories and reach audiences six times faster than the truth [8].The proliferation of fake news has severe implications for societal trust and democratic institutions. For example, recent surveys indicate that two-thirds of EU citizens encounter fake news at least once a week, and over 80% consider it a significant issue for their country and democracy in general [18].

As the challenge of combating fake news becomes increasingly urgent, researchers have turned to artificial intelligence (AI) solutions to address this complex and evolving problem. Initially, traditional machine learning (ML) models were employed to detect fake news by analyzing patterns in text data. These models used various natural language processing (NLP) techniques to preprocess the data before classification. This preprocessing phase often involved tasks such as tokenization (breaking text into smaller units like words or phrases), lemmatization (reducing words to their base forms), and the removal of stop words (common words like "and," "the," or "is," which don't carry significant meaning in this context). The preprocessed data was then passed through various ML algorithms, such as Decision Trees, Support Vector Machines (SVMs), or gradient boosting models like XGBoost, which were tasked with identifying linguistic features, sentiment markers, or other cues that might indicate falsified or biased information. These traditional models offered some success in detecting fake news but often faced challenges in dealing with the complexities and evolving nature of language used in modern misinformation campaigns. For example, they struggled to account for the subtleties of context, irony, and manipulation that are commonly used in fake news stories [24], [27].

In response to these challenges, recent breakthroughs in Large Language Models (LLMs), such as OpenAI's GPT models, have revolutionized the domain of fake news detection. LLMs, which are built with billions of parameters and trained on massive, diverse datasets, have an exceptional ability to understand the intricacies of language and context [7]. Unlike traditional ML models, which rely on manually defined features, LLMs can grasp the broader meaning of text, including contextual nuances, which are crucial in identifying misleading or false information. For example, these models can better detect the subtle ways in which fake news might manipulate facts or appeal to emotions through language that might be overlooked by traditional models.

The key advantage of LLMs is their ability to go beyond simple feature extraction, moving towards a more sophisticated semantic and contextual understanding of news content. This allows them to analyze not just the surface-level linguistic elements, but also the deeper meaning, tone, and intent behind the words. As a result, LLMs have significantly improved the accuracy and robustness of fake news detection, outperforming older ML models in detecting even the most subtle or deceptive forms of misinformation. By focusing on the broader context and understanding how information is framed, LLMs are helping to address the growing problem of fake news in ways that traditional models simply could not [5], [26].

In our paper, we aim to investigate two key questions that remain central to the effectiveness of AI-driven fake news detection:

1. Are LLMs more accurate in detecting fake news than traditional ML algorithms? This question seeks to assess the performance differences between traditional ML algorithms and modern LLMs in identifying fake news. By comparing metrics such as accuracy, precision, recall, and F1-score across various models, we aim to quantify whether the contextual understanding afforded by LLMs translates to better detection rates.

2. Do hybrid models that combine Large Language Models (LLMs) with traditional machine learning models or Small Language Models (SMLs) offer improvements in fake news detection? Hybrid models, which combine the strengths of LLMs with more lightweight or interpretable models, represent an intriguing area of research. These combinations could balance the computational intensity of LLMs with the efficiency and interpretability of smaller models, potentially enhancing overall model performance and providing faster,

more resource-efficient solutions.

II. Related Work

The problem of fake news detection has been extensively studied, with advancements spanning traditional machine learning approaches to the adoption of large language models (LLMs). This section provides a detailed review of relevant works, focusing on specific methodologies and their contributions.

A. LIAR Dataset

1) Fake News Detection Using Machine Learning Approaches: Khanam et al. [12] examined the application of traditional supervised machine learning methods, including Support Vector Machines (SVMs), Random Forest, and XGBoost, for fake news detection. Using TF-IDF for feature extraction, the XGBoost classifier achieved the highest accuracy of 75%. While computationally efficient, the study highlighted the limited scalability of such approaches to multi-modal datasets and their inability to leverage contextual understanding.

2) Fake Detect: A Deep Learning Ensemble Model for Fake News Detection: Aslam et al. [2] proposed a deep learning ensemble model for fake news detection, leveraging the LIAR dataset. The authors employed a hybrid approach using two deep learning models: a Bi-LSTM-GRU network for textual attributes (statements) and a dense model for non-textual attributes (e.g., speaker's job title and context). Preprocessing steps included tokenization, lemmatization, stop-word removal, and word embeddings using FastText.

The proposed model achieved an accuracy of 89.8% and an F1-score of 0.914, outperforming traditional machine learning methods and prior CNN-based approaches. However, challenges included the limited contribution of non-textual attributes to classification and reliance on a single dataset, which limited generalizability to other domains.

3) An Ensemble Machine Learning Approach to Classify Fake News: Hakak et al. [9] developed an ensemble machine learning approach that focused on effective feature extraction. They used the LIAR and ISOT datasets. The study emphasized robust preprocessing techniques, such as tokenization, noise removal, and Named Entity Recognition (NER) to extract features like word count and average sentence length.

The ensemble model combined Decision Tree, Random Forest, and Extra Tree classifiers using a bagging approach for stability. Results demonstrated 100% accuracy on the ISOT dataset and 99.96% training accuracy with 44.15% testing accuracy on the LIAR dataset. However, generalization remained an issue, particularly for the LIAR dataset, due to its complexity and multi-class nature.

4) Fake News Prediction Using Machine Learning Approaches: Mushtaq et al. [17] focused on fake news prediction using the LIAR dataset. Their research employed machine learning classifiers, including Naïve Bayes, Random Forest, Decision Tree, and Neural Networks. Preprocessing steps included data cleaning to remove noise and unnecessary symbols and statistical feature extraction, such as analyzing word distributions and subject categories.

The study highlighted the effectiveness of Naïve Bayes, which achieved a 99% accuracy due to its ability to reduce variance and mitigate overfitting. Compared to other classifiers, Naïve Bayes required less computational time while maintaining high precision and recall. Despite the promising results, challenges included limited exploration of deep learning approaches and reliance on static benchmark datasets, which may not reflect real-time complexities of fake news on social media.

5) A Better Large Language Model Using LoRA for False News Recognition System: Tiwari [23] introduced a framework leveraging Low-Rank Adaptation (LoRA) to fine-tune the LLaMA2-7B language model for fake news detection. LoRA reduces the computational demands of training large-scale language models by decomposing their weight matrices into low-rank components, enabling task-specific adaptation with fewer parameters. The study utilized datasets such as the COVID-19 FakeNews dataset, LIAR dataset, and the FakeNews Challenge dataset. The preprocessing pipeline included text normalization and imbalance correction through class weighting.

The results showed significant performance improvements, with accuracy reaching 97.33% on the COVID-19 dataset and 98.66% on the FakeNews Challenge dataset. These findings underscore the effectiveness of LoRA in optimizing LLMs for resource-constrained environments. Despite these successes, the study noted diminishing returns with extended training durations and emphasized the need for further research on efficient adaptation techniques.

6) A Novel Framework for Fake News Detection Using Double Layer BI-LSTM: Merryton and Augasta [16] developed a novel framework based on Double Layer Bi-LSTM for enhancing fake news detection. By stacking two Bi-LSTM layers, the model captures both short-term and long-term dependencies in textual data. The framework combines traditional preprocessing techniques, such as the Porter Stemmer and TF-IDF vectorization, with deep learning for feature extraction.

The study evaluated the framework on three datasets: the Kaggle Fake_Real_News, LIAR, and Politifact datasets. Results showed a 97.58% accuracy on Kaggle and 83% accuracy on Politifact. However, the model underperformed on the LIAR dataset (61.19%), likely due to its limited ability to handle highly imbalanced classes and nuanced text features. This indicates potential areas for improvement, such as incorporating attention mechanisms for better feature weighting.

7) Fighting Lies with Intelligence: Using Large Language Models and Chain of Thoughts Technique to Combat Fake News: Kareem and Abbas [11] introduced the Chain of Thoughts (CoT) reasoning approach to enhance the interpretability and accuracy of fake news detection systems. By fine-tuning FLAN-T5 and LLaMA-2 with CoT annotations, the models were able to provide logical justifications for their predictions. This method significantly improved the accuracy of binary classification tasks, achieving 84.26% accuracy on a dataset enriched with CoT-annotated records.

While the CoT approach improved transparency, challenges included limited performance gains on multi-class datasets and the need for richer annotation schemes. The study concluded with recommendations for integrating CoT with multimodal data for enhanced generalizability.

8) Re-Search for the Truth: Multi-Round Retrieval-Augmented LLMs for Fake News Detection: Li et al. [14] proposed the STEEL framework, which combines multi-round retrieval mechanisms with LLMs for dynamic evidence collection and claim verification. By sequentially retrieving high-quality evidence until confidence thresholds are met, the framework outperformed traditional single-retrieval methods.

STEEL was evaluated on the LIAR, CHEF, and PolitiFact datasets, achieving F1-Macro scores of 0.714, 0.793, and 0.751, respectively. The multi-round retrieval mechanism significantly improved accuracy in detecting fake news. However, reliance on internet accessibility and the limitations of LLM context lengths were noted as challenges.

9) Fake News Detection with Large Language Models on the LIAR Dataset: Boissonneault and Hensen [4] conducted a detailed evaluation of LLMs like ChatGPT and Google Gemini on the LIAR dataset. Google Gemini achieved an accuracy of 89.4%, outperforming ChatGPT in terms of precision and recall. Despite these strong results, the study highlighted limitations in handling nuanced contextual information, suggesting the need for domain-specific fine-tuning to improve detection of subtle misinformation.

10) A Survey on the Use of Large Language Models (LLMs) in Fake News: Papageorgiou et al. [19] presented a comprehensive survey of LLMs in fake news detection, comparing their performance with traditional machine learning models. The survey highlighted the superior context-aware capabilities of LLMs, which enable them to outperform conventional methods. Challenges such as computational costs, ethical concerns, and susceptibility to biases were discussed, emphasizing the need for future research into more efficient and equitable systems.

B. Other Datasets

1) Integrating Large Language Models and Machine Learning for Fake News Detection: Teo et al. [22] proposed a hybrid method combining LLMs with traditional machine learning algorithms, specifically using ChatGPT-3.5 outputs as features for XGBoost classifiers. This integration leveraged the contextual strengths of LLMs and the efficiency of XGBoost, achieving an accuracy of 96.39%. The study emphasized the potential of hybrid models to balance interpretability and computational efficiency, particularly in resource-constrained settings.

2) CSI: A Hybrid Deep Model for Fake News Detection: Ruchansky et al. [20] presented the CSI framework, which integrates three key elements: content, social context, and user engagement behavior. The model employs a hybrid deep architecture combining LSTM networks for temporal analysis and Singular Value Decomposition (SVD) for user behavior analysis. This approach uniquely captures temporal and user-level patterns, which are often critical for distinguishing fake news from legitimate content.

The study utilized datasets from Twitter and Weibo, achieving an accuracy of 89.2% on Twitter and 95.3% on Weibo. The inclusion of user behavior scores provided additional insights into suspicious activities, enabling a more interpretable classification process. However, the model's reliance on large-scale, annotated user interaction data poses challenges for generalizability across platforms with limited or incomplete engagement data.

3) Bad Actor, Good Advisor: Exploring the Role of Large Language Models in Fake News Detection: Hu et al. [10] proposed the Adaptive Rationale Guidance (ARG) network, which leverages LLMs like GPT-3.5 as advisors rather than decision-makers. ARG integrates LLM-generated rationales into the decision-making process of small language models (SLMs). This hybrid approach achieved 10-15% higher accuracy compared to standalone models while reducing computational demands. However, querying LLMs remains resource-intensive, and selective sampling methods were suggested to mitigate this issue.

4) Adapting Fake News Detection to the Era of Large Language Models: Su et al. [21] explored the challenges of adapting fake news detection systems to a mixed-content landscape dominated by both human-written and machine-generated news. The study investigated biases in existing detectors and proposed strategies to improve robustness by fine-tuning models on datasets augmented with machine-generated content.

The research utilized datasets such as GossipCop++ and PolitiFact++ and augmented them with content generated by ChatGPT. Key findings included the importance of balancing machine-generated content in training data to enhance generalizability. The study also identified biases in popular transformer-based models, such as RoBERTa favoring fake news and ALBERT favoring real news. This work highlighted the need for comprehensive datasets and advanced training strategies to mitigate biases and improve detector reliability.

5) Evaluating the Efficacy of Large Language Models in Detecting Fake News: A Comparative Analysis: Koka et al. [13] conducted a comparative analysis of six LLMs, including GPT-4, Claude, and Mistral, to evaluate their performance in fake news detection. The study employed a balanced dataset of 30 articles and used zero-shot prompting for classification. Results revealed that larger models, such as Claude and GPT-4, achieved near-perfect accuracy and F1 scores, while smaller models like Mistral 7B exhibited higher false-positive rates.

The authors emphasized the superior contextual understanding of larger models but noted that the limited dataset size restricted generalizability. Future work proposed by the authors includes expanding datasets and exploring ensemble methods to integrate outputs from multiple models for improved performance.

6) Large Language Model Agent for Fake News Detection: Li et al. [15] introduced FactAgent, an innovative system combining internal LLM reasoning with external search tools to emulate human expert workflows for fake news detection. The system integrates domain-specific tools and decomposes complex tasks into simpler sub-tasks, achieving notable accuracy gains on datasets such as PolitiFact (88%) and GossipCop (83%).

FactAgent demonstrated flexibility and scalability, particularly in resource-constrained settings. However, the study identified areas for improvement, including the integration of multimodal content and advanced decision-making strategies.

7) Large Language Model-Based Fake News Detection: Aman [1] introduced a lightweight framework combining LLaMA models with task-specific fine-tuning techniques for detecting fake news in text, images, and multimedia formats. The study focused on optimizing LLMs using methods like alignment and mixed-precision quantization, significantly reducing resource usage while maintaining high accuracy. While the framework demonstrated robust performance, real-time application faced challenges due to hardware limitations, emphasizing the need for further optimization for low-resource settings.

8) News Verifiers Showdown: Comparative Performance of LLMs: Caramancion [6] evaluated the performance of prominent LLMs, including GPT-4, Bing AI, Bard, and Claude, on a dataset of 100 fact-checked news articles. GPT-4 achieved the highest score, correctly classifying 71 out of 100 articles, demonstrating its superior contextual analysis capabilities. However, the study highlighted challenges such as hallucinations and false positives, emphasizing the need for model improvements to enhance reliability in real-world applications.

TABLE I
BEST-PERFORMING METHODS AND THEIR ACCURACIES

Best-Performing Method	**Accuracy (%)**
TF-IDF, XGBoost	75
Bi-LSTM-GRU-dense model	89
Ensemble ML Approach	44.15
Naive Bayes	99
LoRA, LLaMA2-7B	62.67
Double Layer Bi-LSTM	97.58
CoT reasoning, FLAN-T5 XXL	84.26
STEEL framework, multi-round evidence retrieval	79.3 (Macro-F1)
Google Gemini	89.4

III. DATASET

Selected over a decade from PolitiFact.com, the publicly available LIAR dataset, developed by Wang [25], consists of 12.8K brief statements classified for truthfulness. Every LIAR record features not just the statement but also a thorough analysis, source references, and metadata including speaker job title, party affiliation, and historical truthfulness. With columns for the statement ID, truthfulness label (e.g., true, largely true, false), statement text, subject, speaker details, and context, the dataset is TSV (tab-separated values).

IV. METHODOLOGY

This section outlines the processes involved in exploring the dataset, performing preprocessing, and conducting feature engineering. The objective is to prepare the data for subsequent stages of model implementation and evaluation. A variety of techniques were applied, including text cleaning, handling missing data, encoding categorical variables, and engineering new features to enhance data quality and improve model performance.

TABLE II
DATASET DISTRIBUTION OVERVIEW

Dataset Split	**Number of Rows**	**Percentage (%)**
Train	10,296	80.0
Test	1,267	9.8
Validation	1,284	10.0
Total	**12,847**	**100.0**

A. Dataset Overview

The dataset used in this study is the LIAR dataset, which consists of labeled statements derived from news articles. To facilitate processing, the data was converted from TSV format to CSV. It includes categorical, numerical, and textual features as outlined below:

- Categorical Features: ”Label,” ”Speaker,” ”Job Title,” ”State,” ”Party,” and ”Context.”
- Numerical Features: ”Barely True Count,” ”False Count,” ”Half True Count,” ”Mostly True Count,” and ”Pants on Fire Count.”
- Textual Feature: ”Statement,” which contains textual information that requires cleaning and transformation before being used in machine learning models.

B. Workflow

The following steps were carried out to prepare the dataset for feature extraction and model training:

1) Data Loading: The dataset was imported into the environment using the pandas library, which provides a structured format for processing.
2) Handling Missing Data: An initial inspection revealed missing values, necessitating imputation and cleaning. Irrelevant columns such as ID were removed, and rows with irrelevance or outliers were excluded. Missing values in categorical columns were replaced with ”Unknown,” while missing numerical values were imputed using their respective median values. Missing values were visualized using a heatmap to ensure proper imputation (see Fig. **??**).
3) Categorizing Political Party: To categorize political parties, a threshold of 200 occurrences was defined. Those featuring lower numbers were categorized as ‘Others’. This in part, aimed at restricting the diversity of party labels and focusing attention on the best represented. A bar graph was generated to show the party affiliation distribution in terms of percentage.
4) Text Cleaning: The ”Statement” column, which contains the primary textual data, underwent a series of cleaning steps. These included:

- Converting text to lowercase
- Removing URLs, numbers, and non-word characters
- Removing extra spaces and stop words
- Lemmatization of words to reduce them to their base forms

This step helped standardize and simplify the text data for further analysis.

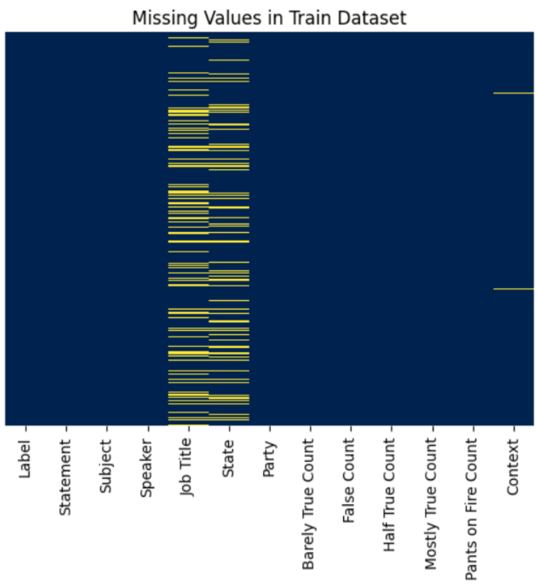

Fig. 1. Missing Values Before Preprocessing

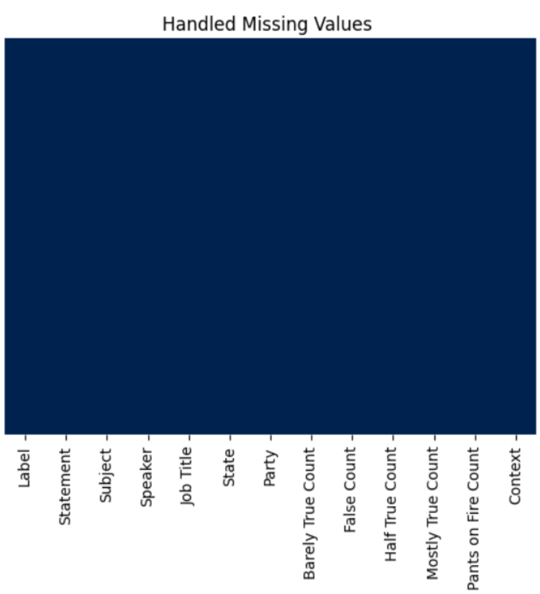

Fig. 2. Missing Values After Preprocessing

5) Feature Engineering: Several new features were engineered to enrich the dataset and enhance the model:
    - Ordinal Encoding: The 'Label' column, which contains categorical truth labels (e.g., "TRUE", "FALSE"), was converted into a numerical format using ordinal encoding, making it compatible with machine learning models.
    - Sentiment Analysis: A sentiment score for each statement was calculated using TextBlob. This feature helps capture the sentiment expressed in the statement, which could be useful for identifying fake news based on emotional tone.
    - False Ratio: A new feature was created to represent the proportion of "False Count" relative to the total of all truth-related counts. This measure captures the relative falsehood in each statement.
6) Data Analysis: After preprocessing, exploratory data analysis (EDA) was conducted to extract insights from the dataset:
    - Label Distribution: A count plot was generated to visualize the distribution of different labels (e.g., "TRUE," "FALSE," "barely-true"), providing insights into potential class imbalances.
    - Feature Correlation: A correlation matrix of numerical features was computed and visualized using a heatmap. This helped identify relationships among features and the detection of multicollinearity.
    - Word Cloud: A word cloud was generated to highlight the most frequently occurring words in the statements, providing a visual representation of the dataset's vocabulary.
    - Sentiment Distribution: A histogram was created to illustrate the distribution of sentiment polarity across the statements, offering insights into the overall emotional tone conveyed in the data.
    - Political Party and Truthfulness: A count plot was used to examine the relationship between political party affiliation and the truthfulness of statements, helping to identify any potential biases related to political affiliations.
    - Speaker Analysis: The ten most frequent speakers were selected, and a count plot was generated to examine the truthfulness of statements attributed to each speaker, providing insights into speaker-specific patterns of accuracy.

References

[1] Mussa Aman. Large language model based fake news detection. In *Soft Computing and Intelligent Systems: Theory and Applications (EUSPN 2023)*, Almaty, Kazakhstan, November 7-9 2023. Kazakh-British Technology University, Elsevier B.V. Available online at www.sciencedirect.com.

[2] Nida Aslam, Irfan Ullah Khan, Farah Salem Alotaibi, Lama Abdulaziz Aldaej, and Asma Khaled Aldubaikil. Fake detect: A deep learning ensemble model for fake news detection. *Complexity*, 2021:1–8, 2021.

[3] BBC Bitesize. What is misinformation?, 2024. Accessed: 2024-11-12.

[4] David Boissonneault and Emily Hensen. Fake news detection with large language models on the liar dataset. *Research Article*, 2024. Creative Commons Attribution 4.0 International License.

[5] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, et al. Language models are few-shot learners. In *Proceedings of NeurIPS 2020*, 2020.

[6] Kevin Matthe Caramancion. News verifiers showdown: A comparative performance evaluation of chatgpt 3.5, chatgpt 4.0, bing ai, and bard in news fact-checking. *University of Wisconsin–Stout Mathematics, Statistics, and Computer Science Department*, 2024. Available at: www.kevincaramancion.com.

[7] Kindra Cooper. Openai gpt-3: Everything you need to know [updated]. https://www.springboard.com/blog/data-science/machine-learning-gpt-3-open-ai/, 2023. Accessed: 2024-12-03.

[8] A. Friggeri, E. E. Garcia, and L. A. González. Study: False news travels faster than true stories on twitter. *MIT News*, 2018. Accessed: 2024-11-12.

[9] Saqib Hakak, Mamoun Alazab, Suleman Khan, Thippa Reddy Gadekallu, Praveen Kumar Reddy Maddikunta, and Wazir Zada Khan. An ensemble machine learning approach through effective feature extraction to classify fake news. *Future Generation Computer Systems*, 117:47–58, 2021.

[10] Beizhe Hu, Qiang Sheng, Juan Cao, Yuhui Shi, Yang Li, Danding Wang, and Peng Qi. Bad actor, good advisor: Exploring the role of large language models in fake news detection. In *Proceedings of the Thirty-Eighth AAAI Conference on Artificial Intelligence (AAAI-24)*, Vancouver, BC, Canada, 2024. Association for the Advancement of Artificial Intelligence. CAS Key Lab of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences; University of Chinese Academy of Sciences; National University of Singapore.

[11] Waleed Kareem and Noorhan Abbas. Fighting lies with intelligence: Using large language models and chain of thoughts technique to combat fake news. In *Proceedings of the 43rd Annual International Conference on Artificial Intelligence (SGAI)*, 2023.

[12] Z Khanam, B N Alwasel, H Sirafi, and M Rashid. Fake news detection using machine learning approaches. *IOP Conference Series: Materials Science and Engineering*, 1099:012040, 2021. International Conference on Applied Scientific Computational Intelligence using Data Science (ASCI 2020), 22nd-23rd December 2020, Jaipur, India.

[13] Sahas Koka, Anthony Vuong, and Anish Kataria. Evaluating the efficacy of large language models in detecting fake news: A comparative analysis. *Preprint*, 2024.

[14] Guanghua Li, Wensheng Lu, Wei Zhang, et al. Re-search for the truth: Multi-round retrieval-augmented large language models are strong fake news detectors. *Preprint arXiv*, 2024.

[15] Xinyi Li, Yongfeng Zhang, and Edward C. Malthouse. Large language model agent for fake news detection. *Preprint arXiv*, 2024.

[16] Adline Rajasenah Merryton and M. Gethsiyal Augasta. A novel framework for fake news detection using double layer bi-lstm. In *Proceedings of the 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, 2023.

[17] Ansa Mushtaq, Muhammad Javaid Iqbal, Saba Ramzan, Sobia Yaqoob, Ali Asif, and Inam Ul Haq. Fake news prediction using machine learning approaches. *Bulletin of Business and Economics*, 13(1):683–689, 2023.

[18] Council of Europe. Dealing with propaganda, misinformation and fake news, 2024. Accessed: 2024-11-12.

[19] Eleftheria Papageorgiou, Christos Chronis, Iraklis Varlamis, and Yassine Himeur. A survey on the use of large language models (llms) in fake news. *Department of Informatics and Telematics, Harokopio University of Athens*, 2024. Correspondence: varlamis@hua.gr.

[20] Natali Ruchansky, Sungyong Seo, and Yan Liu. Csi: A hybrid deep model for fake news detection. In *Proceedings of the 26th ACM International Conference on Information and Knowledge Management (CIKM)*, 2017.

[21] Jinyan Su, Claire Cardie, and Preslav Nakov. Adapting fake news detection to the era of large language models. *Preprint arXiv*, 2024.

[22] Ting Wei Teo, Hui Na Chua, Muhammed Basheer Jasser, and Richard T.K. Wong. Integrating large language models and machine learning for fake news detection. In *2024 20th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)*, Langkawi, Malaysia, March 1-2 2024. Department of Computing and Information Systems, Sunway University, IEEE.

[23] Saransh Vinodchandra Tiwari. A better large language model using lora for false news recognition system. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*, 12(6), 2024.

[24] S. Vosoughi, D. Roy, and S. Aral. The spread of true and false news online. *Science*, 359(6380):1146–1151, 2018.

[25] William Yang Wang. ”liar, liar pants on fire”: A new benchmark dataset for fake news detection. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (ACL 2017), Short Papers*, pages 422–426, Vancouver, BC, Canada, July 2017. Association for Computational Linguistics.

[26] R. Zellers, A. Holtzman, L. Wu, M. Dohan, and Y. Choi. Defending against neural fake news. In *Proceedings of NeurIPS 2019*, 2019.

[27] Z. Zhao, P. Resnick, and Q. Mei. Enquiring minds: Early detection of fake news. In *Proceedings of the 24th International Conference on World Wide Web*, 2015.