



Comparative Evaluation of Lightweight Pre-trained Language Models for Fake News Detection

A Research Proposal Submitted in Partial Fulfillment of the
Requirements for the Course: Scientific Research

Ameed Othman, Karim Mithqal, Waleed Dweikat

Supervisor:
Dr. Ahmed Alia

An-Najah National University
Faculty of Information Technology and Artificial Intelligence
Department of Computer Science

May 4, 2025

1 Introduction

In this digital age, especially with the rise of large language models (LLMs), it is very easy to share misinformation online. Anyone with a digital device can use an AI chatbot to generate fake content that is very well written and sounds convincing. These LLMs can even imitate established and renowned authors in their writing style and format. Therefore, it's essential to develop modern methods for fake news detection (FND). Recent research shows that deep learning techniques promise great potential for FND, often outperforming traditional machine learning methods. The problem with these DL-based models is that they are very resource intensive and require great compute power. To this end, we propose a comparative study, in which we compare light-weight pre-trained language models to check how well they can detect misinformation in resource-constrained environments. These are the main challenges that we will face in our research:

1. We might be trading accuracy for less compute power, and thus the models might not be able to effectively detect misinformation.
2. Using smaller datasets might lead models not to generalize well on new types of misinformation, especially those related to more recent news in 2025.

The next section will review the most efficient techniques for fake news detection.

2 Brief Literature Review

In recent years, the spread of misinformation online is a major problem, especially with the intelligence and modernity of LLMs that can generate very identical fake content[5]. Any one can create fake news as long as there is a digital device and some kind of AI tool involved coupled with a bit of knowledge, hence the need for robust modern methods for fake news detection (FND) [2]. Traditional machine learning methods are disappointing, they are limited and complex in handling online content, while deep learning methods are more efficient with astonishing performance [9]. The existing FND approaches can be generally classified into three main categories: traditional machine learning methods, deep learning based methods, and hybrid approaches [8]. Traditional machine learning methods rely on manual features labeling such as text length, readability score , lexical diversity [4]. These features are used to train classifiers like support vector machines or random forests [13]. Due to challenges such as nuanced semantic patterns of modern misinformation , these approaches fail to FND accurately [7]. On the other hand, deep learning based methods have shown superior performance by automating the process of extracting features from text [10]. Convolutional neural networks (CNNs) have been doing a good job in capturing local semantic patterns, whereas recurrent neural networks (RNNs) and their variants are proficient at modeling sequential dependencies in text [1]. More recently, machine learning models like have achieved remarkable results in hybrid approaches [3]. The literature FND highlights several trends and challenges. First, while deep learning models outperform traditional methods, they require great compute power , resources , and large data sets [9]. Second the interpretability of these models remains a critical concern, especially in understanding the decision-making process is as important as the prediction itself [12]. In our proposed study, we want to address these challenges by evaluating lightweight pre-trained language models as efficient alternatives for FND in resource-constrained environments. These models offer a good balance between computation efficiency and detection accuracy [6].

3 Overall Goal and Research Questions

This project aims to conduct an empirical comparison among several lightweight pre-trained language models for fake news detection in resource-constrained environments. The goal is to identify which model achieves the best balance between detection accuracy (e.g., F1-score) and resource efficiency (e.g., inference time, memory usage), enabling practical deployment on limited hardware.

1. Among the lightweight pre-trained language models, which one achieves the best performance in fake news detection while using minimal computational resources?
2. Can these lightweight models maintain competitive performance when trained on limited or imbalanced datasets?

4 Methodology

In this research we use a **quantitative approach** to evaluate lightweight pre-trained language models for fake news detection. The main steps are as follows:

1. **Collecting and preparing a Dataset:** We have decided to use LIAR2[11], which builds on the original LIAR dataset that has been used often in FND research. This step involves downloading the datasets from reliable sources, conducting exploratory data analysis, and performing data wrangling and cleaning to prepare the dataset for consumption by the models. We will use LIAR2 for fine-tuning and we will gather real world news articles as well as AI-generated ones for testing generalization.
2. **Selecting pre-trained lightweight models:** In this step, we will select which models to experiment with and evaluate. Since we are working with limited resources, we will focus on very lightweight models like DistillBERT.
3. **Fine-Tuning the models:** In this step, we will fine-tune each selected model on the LIAR2 dataset, optimizing both for performance and resource efficiency.
4. **Evaluation:**
 - (a) **Analyzing resource consumption:** In this step, we will analyze how each model consumes resources like CPU and memory during inference.
 - (b) **Evaluating performance:** We will also evaluate the detection performance of each model, measuring metrics like F1-score and accuracy.
5. **Comparing evaluation:** In this step, we will compare results from both evaluations across different models to identify the best performance-resource tradeoffs. Based on these findings, we may return to step 3 to further optimize promising models.
6. **Generalization testing:** In this final step, we will test the most promising models on recent news data as well as fake AI-generated news data to evaluate real-world effectiveness.

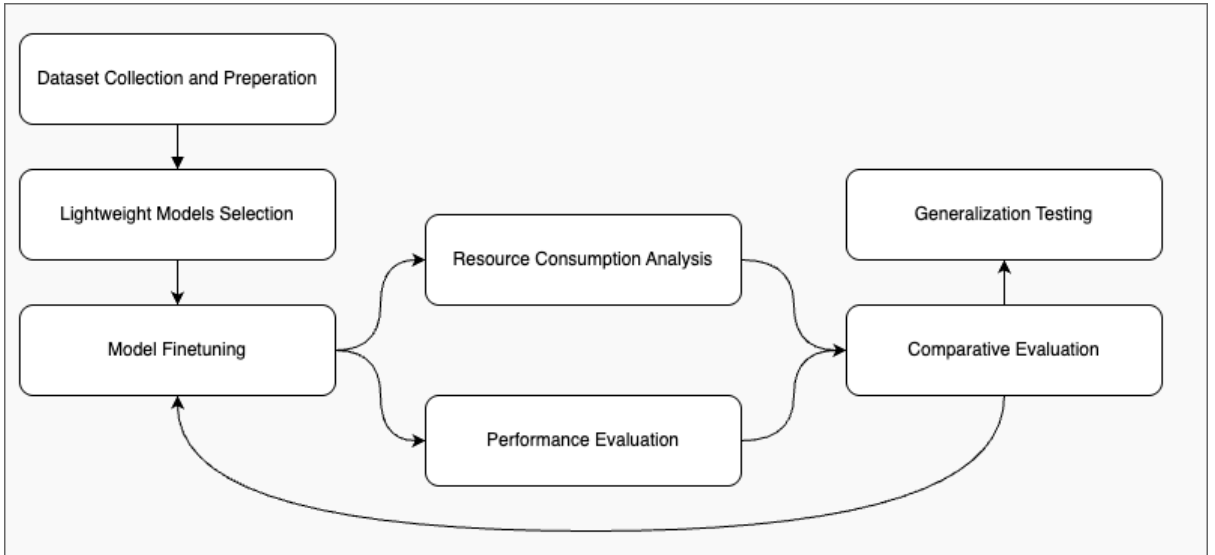


Figure 1: Research methodology flow chart.

5 Timeline

Date range	Work
02.03-11.03	Idea exploration
12.03-19.03	Establishing shared Zotero library and collecting papers
20.03-29.03	Writing the research proposal
30.03-02.04	Eid Al-Fitr holiday
03.04-06.04	Prepare and deliver presentation of research proposal
07.04-13.04	Dataset collection and preparation
14.04-21.04	Models selection
22.04-05.05	Model fine-tuning
06.05-12.05	Model evaluation
13.05-19.05	Comparative analysis and model refinement
20.05-26.05	Generalization testing
27.05-02.06	Writing final documentation and presenting work

Table 1: Timeline

6 Author Contributions

- **Waleed:** uploaded An-Najah University logo.
- **Ameed:** transformed the title page into LaTeX.
- **Ameed:** transformed the Introduction section into LaTeX.
- **Ameed:** changed the titles of the sections to match our research proposal and added the Author Contributions section.
- **Ameed:** transformed the methodology section into LaTeX.
- **Ameed:** uploaded research methodology flow chart and inserted it into the document.
- **Ameed:** created timeline table
- **Karim:** transformed the Literature Review section into LaTeX.
- **Karim:** transformed the References section into LaTeX.
- **All members:** Collaborated to revise the overall goal and objectives.
- **Ameed & Karim:** organized pages flow and added \newpage where necessary.
- **Waleed:** transformed Goal and Research Questions section into LaTeX and made it clear that we aim to do empirical comparisons.

References

- [1] Mohammed Al-Alshaqi, Danda B Rawat, and Chunmei Liu. Ensemble techniques for robust fake news detection: Integrating transformers, natural language processing, and machine learning. *Sensors*, 24(18):6062, 2024.
- [2] Jawaher Alghamdi, Yuqing Lin, and Suhuai Luo. Unveiling the hidden patterns: A novel semantic deep learning approach to fake news detection on social media. *Engineering Applications of Artificial Intelligence*, 137:109240, 2024.
- [3] Beatriz Flãmia Azevedo, Ana Maria AC Rocha, and Ana I Pereira. Hybrid approaches to optimization and machine learning methods: a systematic literature review. *Machine Learning*, 113(7):4055–4097, 2024.
- [4] Mohammed E. Almandouh, Mohammed F Alrahmawy, Mohamed Eisa, Mohamed Elhoseny, and AS Tolba. Ensemble based high performance deep learning models for fake news detection. *Scientific Reports*, 14(1):26591, 2024.
- [5] Linmei Hu, Siqi Wei, Ziwang Zhao, and Bin Wu. Deep learning for fake news detection: A comprehensive survey. *AI open*, 3:133–155, 2022.
- [6] Ajay Kumar and James W Taylor. Feature importance in the age of explainable ai: Case study of detecting fake news & misinformation via a multi-modal framework. *European Journal of Operational Research*, 317(2):401–413, 2024.
- [7] Despoina Mouratidis, Andreas Kanavos, and Katia Kermanidis. From misinformation to insight: Machine learning strategies for fake news detection. *Information (2078-2489)*, 16(3), 2025.
- [8] Konstantinos I Roumeliotis, Nikolaos D Tselikas, and Dimitrios K Nasiopoulos. Fake news detection and classification: A comparative study of convolutional neural networks, large language models, and natural language processing models. *Future Internet*, 17(1), 2025.
- [9] Kayato Soga, Soh Yoshida, and Mitsuji Muneyasu. Exploiting stance similarity and graph neural networks for fake news detection. *Pattern Recognition Letters*, 177:26–32, 2024.
- [10] Aswini Thota, Priyanka Tilak, Simrat Ahluwalia, and Nibrat Lohia. Fake news detection: a deep learning approach. *SMU Data Science Review*, 1(3):10, 2018.
- [11] Cheng Xu and M-Tahar Kechadi. An enhanced fake news detection system with fuzzy deep learning. *IEEE Access*, 12:88006–88021, 2024.
- [12] Junxiao Xue, Yabo Wang, Yichen Tian, Yafei Li, Lei Shi, and Lin Wei. Detecting fake news by exploring the consistency of multimodal data. *Information Processing & Management*, 58(5):102610, 2021.
- [13] Abu Sarwar Zamani, Aisha Hassan Abdalla Hashim, Sara Saadeldeen Ibrahim Mohamed, and Md Nasre Alam. Optimized deep learning techniques to identify rumors and fake news in online social networks. *Journal of Computational and Cognitive Engineering*, 2022.