



IJRASET

International Journal For Research in
Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** VI **Month of publication:** June 2024

DOI: <https://doi.org/10.22214/ijraset.2024.63069>

www.ijraset.com

Call:  08813907089

E-mail ID: ijraset@gmail.com

A Better Large Language Model Using Lora for False News Recognition System

Saransh Vinodchandra Tiwari

ASM Institute of Management and Computer Studies, University of Mumbai, C-4, Wagle Industrial Estate, Near Mulund Check Naka, Thane West, Opp. Aplab, Mumbai - 4000604, Maharashtra, India.

Abstract: Fake news detection is an important field of study in Natural Language Processing (NLP); yet, problems still exist since there are few well-labeled datasets for AI models to be trained on, and models' efficacy is impacted by privacy and ethical issues. Developing solid algorithms is further complicated by the disparity between false and real news. Previous techniques need to take into account the capacity to take into account both the unique conditions in which the material is provided and the larger narrative in order to handle the difficulty of contextual knowledge in spotting false news. The study investigates the possibilities of using big language models—LLaMA2-7B in particular—for the identification of false news. The study outlines three goals: 1) creating a benchmark dataset that is customized for use in a fake news dataset; 2) using the Language Model (LLM) to develop a fake news detection system through painstakingly fine-tuned training; and 3) assessing the performance of the final model. Using the LLaMA2-7B model, gathering and analyzing datasets of actual and fake news, training the LLM with a variety of real news, and assessing the model's predictive power of news authenticity are all part of the technique. In spite of limited computing resources, the LLaMA2-7B model studies yielded an impressive 96.66% accuracy rate in detecting false news, outperforming earlier technologies and demonstrating the effectiveness of large language models in distinguishing different kinds of fake news. The results imply that using such models has potential for raising the detection of fake news and offering a powerful defense against false information.

I. INTRODUCTION

In the field of Natural Language Processing (NLP), fake news detection has become an important task. This importance has increased throughout the 2010s, especially during the COVID-19 pandemic. Even while the epidemic served as a sobering reminder of the terrible effects that false information may have, the significance of spotting fake news continues even as the pandemic fades, with focus now moving beyond the realm of false information connected to COVID-19.

However, a number of issues have made it more difficult for fake news detecting systems to be effective. The primary factor limiting the availability of high-quality labeled datasets for AI model training is their scarcity, which is further exacerbated by ethical and privacy concerns. Furthermore, the disparity between authentic and fraudulent news makes it more difficult to create reliable detection methods.

Current methods must change in order to overcome these obstacles and improve the contextual knowledge needed for reliable fake news detection. In particular, it is necessary to have the ability to take into account both the particular context in which the information is delivered and the larger story.

This work investigates the possibilities of using big language models—LLaMA2-7B in particular—for the goal of identifying false news. The study aims to accomplish three goals: first, create a benchmark dataset specifically designed for identifying false news; second, use the Language Model (LLM) to meticulously fine-tune training to create a detection system; and third, assess the effectiveness of the final model.

The technique comprises the following steps: the LLaMA2-7B model is deployed; datasets containing true and fake news articles are acquired and processed; the LLM is trained using a variety of real news samples; and the model's accuracy in predicting the veracity of news items is evaluated.

Large language models are useful in distinguishing different kinds of fake news, as demonstrated by the studies carried out with the LLaMA2-7B model, which produced an impressive accuracy rate of 96.66% in fake news identification, surpassing earlier technologies. This result was achieved despite computing constraints.

These findings have important ramifications since they indicate that using these models could increase the detection of fake news and offer a useful tool for countering false information in public discourse.

II. PROBLEM STATEMENT

A. Difficulties with AI-based Identifying False News

The detection of fake news is difficult because of imbalanced datasets, language manipulation, adversarial tactics, contextual understanding requirements, and data scarcity. AI models find it challenging to distinguish between accurate and inaccurate information because of its intricacy.

B. Previous Approaches and Their Drawbacks

Previous approaches with strong loss functions and adaptive training included BERT and RoBERTa. Nevertheless, despite their prowess in text generation and categorization, large language models (LLMs) like GPT-3.5 struggle with subtle manipulations and don't have a comprehensive method for identifying false news.

C. Using LoRA and LLMs for Innovation

Initially, there were difficulties using LangChain to run LLMs since there was insufficient indirect evidence. Nevertheless, these models could more accurately determine the veracity of news by pre-training LLMs with LoRA on large text corpora and optimizing the mon annotated fake news datasets.

D. Developments in the Detection of Fake News

The study presents a novel method that reduces the requirement for ongoing data gathering and retraining by employing LLMstological inference to determine the veracity of news. Additionally, it looks into teaching LLMs to focus on particular subjects, improving their capacity to confirm the truth of news.

III. LITERATURE REVIEW

One important usage of natural language processing (NLP) approaches is the detection of false news. This work will employ open-source large language models such as LLaMA2 or ChatGLM2 for the purpose of labeling and collecting news stories in order to produce datasets for data analysis.

Because of its effective text training and classification capabilities via Hugging Face interfaces, practical studies showed that LLaMA2 is more suitable for classifying fake news, whereas ChatGLM2 lacks the functionality and documentation required for this experiment.

Unlike the limited information available for GPT-4, LLaMA2, an advanced open-source model by MetaAI, stands out for its detailed documentation and inclusivity, with up to 70 billion parameters and enhanced token count and context length. It provides invaluable resources for researchers in the large language model domain.

Introduced in 2017, the Transformer is a deep learning model that is essential to natural language processing because of its self-attention mechanism, which enables sequential input to be processed in parallel, greatly enhancing tasks like text production and language translation.

The Bert language model was developed as a result of the shift from basic machine learning algorithms like Random Forest, SVM, and Naive Bayes to more sophisticated NLP technologies like RNN, CNN, LSTM, and Vaswani's Transform algorithm.

Large conversational language models such as ChatGPT have emerged and become significant, bringing to light the substantial resources needed for their training as well as the phases of pretraining, fine-tuning, and RLHF. The direct application of news authenticity checks using the Transformers library to go around LangChain, and the usage of LoRA technology to incorporate big language models such as LLaMa2 or ChatGLM2 for logical reasoning in fake news detection.

IV. RESEARCH METHODOLOGY

In our work, we use the Transformers library for sequence classification tasks to fine-tune the pre-trained LLaMA2, a Large Language Model based on the Transformer architecture with self-attention layers.

This model efficiently ascertains the validity of news, having been trained on both genuine and fraudulent news from a dataset via the Transformers interface.

Using a Large Language Model (LLM), processing datasets of true and false news, training the LLM with a variety of real news, and assessing the algorithm's capacity to discern between true and false news are all part of this research approach for fake news identification.

Using sources like Science Direct and Scopus, the study offers a thorough literature assessment on false news detection, emphasizing the difficulties, approaches, and developments in the field. It also highlights the significance of the LoRA technique in the Large Language Model approach.

In order to improve the model's accuracy in identifying fake news and handle issues like low-quality false news and adjusting to the changing nature of fake news, the study trains the model using a range of news data, including a specially created database.

$$y = Wx$$

Let x be the input vector and W be the weight matrix for the LLaMa model. The output of the model is given by the following formula.

$$L(W_s) = \sum_{i=1}^N \left(y_t^{(i)} - y_s^{(i)} \right)^2$$

Let W_t and W_s be the weight matrices for the teacher and student models, respectively. The student model minimizes the following loss function during training, so for the t th input example, the teacher and student models' outputs are and, respectively.

Algorithm: LoRA

- 1: Input: LLM_weights
- 2: Output: Adapted_model
- 3:
- 4:
- 5: # Step 1: Decompose LLM's weights
- 6: $(U, A) = \text{decompose}(\text{LLM_weights})$
- 7:
- 8: # Step 2: Freeze low-rank matrix U and update adaptation parameters A
- 9: repeat until model converges:
- 10: $U = \text{freeze}(U)$
- 11: $A = \text{update_parameters}(A)$ # Update A to improve performance on the target task
- 12:
- 13: # Output the adapted model
- 14: $\text{Adapted_model} = \text{combine}(U, A)$

V. RESULT AND DISCUSSION

Summarization of the experimental Result

| Model score | LIAR dataset | COVID-19 FakeNews Dataset | FakeNews Challenge dataset |
|---------------|--------------|---------------------------|----------------------------|
| LLaMA2 Score | 62.67% | 97.33% | 98.66% |
| SVM Score | - | 91.46%[60] | - |
| Decision Tree | 58.66% | - | 99.55%[61] |
| KnowBert-W+W | 28.95%[45] | 31.46%[60] | - |
| Naive Bayes | - | - | 93.73%[61] |

Though it lost out to Decision Tree in one instance, LLaMA2, improved by LoRA technology, outperformed conventional models like KnowBert-W+W and SVM in fake news detection after 12 training rounds. This suggests that, given specialized, high-quality data, LLaMA2 has the potential to be a domain-specific fake news identifier.

VI. RESULT

The LoRA technology was used to train and fine-tune LLaMA2-6B, which showed remarkable prediction accuracy in identifying false news across three datasets: 62.67% for LIAR, 97.33% for COVID-19, and 98.66% for the FakeNews Challenge Dataset.

Through rigorous fine-tuning, our results demonstrate the efficacy of LoRA-based Large Language Models, such as LLaMA2, in the detection of fake news, demonstrating their capability in domain-specific tasks.

Although LLaMA2 is effective, its high CPU, GPU, and memory resource demand poses difficulties for its adoption because it reduces speed and concurrency and requires significant computational expenses.

VII. DISCUSSION

While resource demands on concurrency limit its effectiveness and speed, the LLAMA2 large language model offers a more flexible approach to fake news detection than traditional methods. During training, it requires significant GPU, CPU, and memory resources, but less so after deployment.

VIII. CONCLUSION

The research effectively showcased the efficacy of the LLaMA2 big language model in discerning the veracity of news, with noteworthy prediction accuracy rates (63.98% and 95.33%) that exceeded the performance of prior algorithms. With a high accuracy rate of 96%, LLaMA2 demonstrated its effectiveness in identifying fake news, meeting early expectations and representing a significant advancement in the field of fake news detection technology. According to the research, more thorough and effective fake news detection solutions can be achieved by fusing the advantages of LLM-based techniques like LLaMA2 with the capabilities of conventional false news detection systems.

The potential and limitations of large language models in this sector are highlighted by the fact that LLaMA2, a large language model, outperformed Decision Tree (99.55%) in fake news detection, with a greater accuracy (96%), compared to the Naive Bayes method (93.73%).

Large language models have unique properties that defy common notions about training efficacy. Increasing the number of training sessions using LLaMA did not consistently improve prediction accuracy, contrary to standard machine learning predictions.

The fact that the ideal number of training sessions for LLaMA varies from the second to the eleventh session indicates that evaluations of training methodologies for large language models are necessary, and open-ended questions should be left open for further investigation.

REFERENCES

- [1] <https://www.researchgate.net/publication/354079121> Thai Fake News Detection Based on Information Retrieval Natural Language Processing and Machine Learning.
- [2] <https://www.researchgate.net/publication/270594887> A weighted LS-SVM based learning system for time series forecasting
- [3] <https://www.researchgate.net/publication/223307915> An extension of the naive Bayesian classifier
- [4] <https://www.researchgate.net/publication/220543019> Improved Use of Continuous Attributes in C45
- [5] <https://www.researchgate.net/publication/378696736> AN IMPROVED LARGE LANGUAGE MODEL FOR FAKE NEWS RECOGNITION SYSTEM
- [6] Neural network design, Jan 2014, H B Demuth, H. B. Demuth, (2014). Neural network design.
- [7] Aligning Language Models to Follow Instructions, Jan 2022, Openai, OpenAI. "Aligning Language Models to Follow Instructions". 2022. [Online]. Available: <https://openai.com/blog/instruction-following/> (2022).
- [8] Fake News Detection, 2022, Sameer Patel, Sameer Patel. "Fake News Detection". Kaggle. 2022. [Online]. Available: <https://www.kaggle.com/code/therealsampat/fake-news-detection>
- [9] Langchain, Accessed: 2022, A I Langchain, Langchain AI. "Langchain". GitHub. Accessed: 2022. [Online]. Available: <https://github.com/langchain-ai/langchain>
- [10] LLaMA: Open and Efficient Foundation Language Models, Dec 2022, Z Dai, Z Wang, X Chen, Y Sun, P Zhang, X Hu, Y Li, J Devlin, M.-W Chang, K Lee, Dai, Z., Wang, Z., Chen, X., Sun, Y., Zhang, P., Hu, X., Li, Y., Devlin, J., Chang, M.-W., & Lee, K. (2022, December). LLaMA: Open and Efficient Foundation Language Models. In Proceedings of the 36th Conference on Neural Information Processing Systems (pp. 19173-19184). NeurIPS.
- [11] Low-Rank Adaptation of Large Language Models, Jun 2021, E J Hu, Y Shen, P J Wallis, Hu, E. J., Shen, Y., Wallis, P. J., et al. (2021, June). Low-Rank Adaptation of Large Language Models. arXiv preprint arXiv:2106.09685.



10.22214/IJRASET



45.98



IMPACT FACTOR:
7.129



IMPACT FACTOR:
7.429



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24*7 Support on Whatsapp)