F² (FutureFiction): Detection of Fake News on Futuristic Technology

MSVPJ Sathvik
Raickers AI
Hyderabad
Telangana, India
msvpjsathvik@gmail.com

Venkatesh Velugubantla
Meridian Cooperative
Atlanta
Georgia, USA
venki.v@gmail.com

Ravi Teja Potla
Slalom
Houston
Texas, USA
raviteja.potla@gmail.com

Abstract

There is widespread of misinformation on futuristic technology and society. To accurately detect such news, the algorithms require up-to-date knowledge. The Large Language Models excel in the NLP but cannot retrieve the ongoing events or innovations. For example, GPT and it's variants are restricted till the knowledge of 2021. We introduce a new methodology for the identification of fake news pertaining to futuristic technology and society. Leveraging the power of Google Knowledge, we enhance the capabilities of the GPT-3.5 language model, thereby elevating its performance in the detection of misinformation. The proposed framework exhibits superior efficacy compared to established baselines with the accuracy of 81.04%. Moreover, we propose a novel dataset consisting of fake news in three languages English, Telugu and Tenglish of around 21000 from various sources.

1 Introduction

In the rapidly evolving landscape of futuristic technology, misinformation has become a pervasive and concerning issue. As groundbreaking innovations such as artificial intelligence, quantum computing, and advanced robotics continue to shape the future, the spread of inaccurate or exaggerated information about these technologies can have profound effects(Marche et al., 2023). Misinformation can distort public perceptions, creating unwarranted fears or unrealistic expectations about the capabilities and implications of these technologies. This, in turn, may lead to misguided policy decisions, hinder the adoption of beneficial technologies, or even fuel unnecessary public concerns that impede the responsible development of emerging innovations(Raponi et al., 2022; Wang et al., 2023).

Moreover, misinformation in the realm of futuristic technology can contribute to a lack of trust in scientific advancements and technological progress. When individuals are exposed to sensationalized or

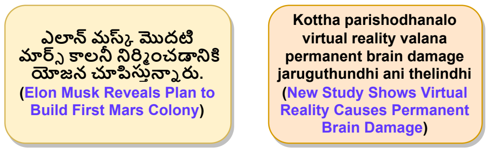

Figure 1: Examples of the fake news related to futuristic technology and society in Telugu and Tenglish. The text in blue color is the translation of the news in English.

inaccurate portrayals of futuristic technologies, it erodes the foundation of public confidence in the scientific community and the technology industry. This erosion of trust can impede collaboration between researchers, policymakers, and the public, hindering the collective efforts needed to navigate the ethical, social, and economic challenges associated with emerging technologies. To address this issue, it is crucial for scientists, technologists, and communicators to prioritize accurate and accessible information, fostering a more informed and discerning public that can engage with the future of technology in a responsible and constructive manner.

To effectively identify and counteract such disinformation, detection algorithms need real-time data to understand the context and verify the accuracy of the information being circulated. Timely updates ensure that the models can recognize and respond to emerging trends, preventing the amplification of false narratives that may contribute to confusion, panic, or even influence public opinion and policy decisions during critical moments in a conflict. To facilitate this up-to-time knowledge update, ontologies and graphs can be employed, offering a structured representation of information that aids in discerning patterns and relationships(Xue and Liu, 2023; Xie et al., 2023).

However, the construction of such ontologies and graphs is a meticulous process that demands time and expertise. Additionally, the rigidity of these structures makes them less adaptable when

Proceedings of the Fifth Workshop on Speech, Vision, and Language Technologies for Dravidian Languages, pages 1–9

transitioning to different domains or subjects. The intricate task of preparing and maintaining these knowledge structures poses a challenge to swiftly respond to evolving scenarios or to seamlessly shift focus to other areas of concern. To overcome these challenges, a pragmatic approach involves leveraging Google's extensive and constantly updated knowledge base. Google serves as a reservoir of real-time information on a myriad of topics, including geopolitical events and war-related developments. By tapping into this vast repository, we can circumvent the time-consuming process of manual ontology creation and instead harness the immediacy and breadth of Google's knowledge.

By integrating Google's dynamic knowledge with the natural language processing (NLP) capabilities of GPT (Generative Pre-trained Transformer), we create a potent synergy. GPT's proficiency in understanding and generating human-like text, coupled with the real-time insights provided by Google, empowers the system to make more informed and timely predictions regarding the authenticity or falsity of information related to war.

This fusion of GPT's linguistic prowess with Google's up-to-date knowledge not only enhances the accuracy of fake news detection but also ensures adaptability to the ever-evolving landscape of information. As a result, this approach not only improves the predictive capabilities in the context of war-related news but also establishes a robust framework that can be extended to different domains, demonstrating a versatility that is crucial in the fast-paced and diverse world of information analysis.

How can we seamlessly integrate Google's knowledge into GPT? One approach involves leveraging Langchain, or alternatively, employing prompting techniques. However, it's crucial to note that these techniques are essentially prompts and might not outperform, especially since GPT isn't explicitly trained to detect fake news. Addressing this necessitates additional training within the context. In this paper, we present a method on how to effectively infuse GPT with knowledge derived from Google, enhancing its capabilities.

The key contributions of our work is as follows:

1. **Novel dataset:** We present a novel dataset with gold human labelled dataset in three languages, Telugu, English and Tenglish.

2. We have implemented baselines on latest approaches like Langchain, GPT-3.5, etc.

3. **Novel Approach:** We present a new algorithm by leveraging Google's knowledge and GPT's capabilities.

2 Related Work

Several approaches have been proposed for detecting and mitigating the spread of fake news across diverse linguistic and thematic domains. Schütz (2023) introduced a disinformation detection method that leverages knowledge infusion through transfer learning and visualizations. Rehm et al. (2018) presented an infrastructure for handling fake news and online media phenomena, incorporating both automatic and manual web annotations. Zhu et al. (2022) proposed a memory-guided multi-view multi-domain fake news detection framework, emphasizing the importance of multi-modal information. Duong et al. (2023) utilized knowledge graph, Datalog, and KG-BERT for fact-checking Vietnamese information.

Ahmed et al. (2022) focused on automatically generating temporally labeled data using positional lexicon expansion for the purpose of estimating the focus time of news articles. Singhal et al. (2022) established FactDrill, a data repository containing fact-checked social media content, facilitating the study of fake news incidents in India. Thaokar et al. (2022) developed a multi-linguistic fake news detector for Hindi, Marathi, and Telugu, emphasizing the importance of linguistic diversity in detection models.

Raja et al. (2023) proposed a method for fake news detection in Dravidian languages using transfer learning with adaptive fine-tuning, addressing linguistic nuances. Yigezu et al. (2023) explored abusive comment detection in Dravidian languages, employing a deep learning approach. Briskilal et al. (2023) introduced an ensemble method for classifying Telugu idiomatic sentences using deep learning models, contributing to the understanding of local linguistic patterns.

Arya et al. (2022) leveraged question answering to understand context-specific patterns in fact-checked articles in the global South. Ren et al. (2023) proposed fake news classification using tensor decomposition and a graph convolutional network. Xie et al. (2023) introduced a knowledge graph-enhanced heterogeneous graph neural network for fake news detection, emphasizing the importance of structured information. Che et al. (2023) proposed tensor factorization with sparse and graph regularization for fake news detection on social networks.

Han et al. (2021) discussed the generation of fake documents using probabilistic logic graphs, providing insights into potential adversarial techniques. Ding et al. (2022) introduced Metadetector, a meta-event knowledge transfer approach for fake news detection. Zhu et al. (2021) presented a knowledge-enhanced approach for fact-checking and verification, highlighting the role of knowledge graphs. Clark et al. (2021) integrated transformers and knowledge graphs for Twitter stance detection, demonstrating the effectiveness of combining these two powerful techniques.

Our proposed dataset focuses specifically on fake news related to futuristic technology and society, providing a unique thematic perspective. Moreover, our algorithm incorporates a fusion of Google's knowledge and the GPT-3.5 model, offering a novel and robust approach to fake news detection in this distinctive domain. This combination of thematic focus and advanced model integration contributes to the enrichment and diversification of the existing landscape of fake news detection methodologies.

3 Data

Data is sourced from Twitter posts and news articles, with newspapers such as The Hindu, Eenadu, Deccan Chronicle, Sakshi, Andhrajyothi, Times of India, and The Indian Express contributing to the dataset. To uphold anonymity and adhere to ethical considerations, the information collected from both newspapers and social media posts is paraphrased. For the paraphrasing of English and Telugu data, a freely available paraphrase tool (paraphrase-tool.com), accommodating multiple languages, is employed. Specifically for Tenglish data, annotators are tasked with manual paraphrasing. The collected data pertains to three languages: Telugu, English, and Tenglish, all focusing on futuristic technology and society. All the news articles and posts gathered are till the May 2023.

Data Annotation: Our goal was to acquire manual ground-truth labels indicating the presence of a string evidence to claim the information is fake or real. We distributed the collected data in batches to annotators, ensuring that each data point was assessed by multiple annotators to minimize labeling errors. Additionally, we ensured that the same annotator did not review the same pairs across batches.

Table 1: Statistics of the Dataset

Source	label 0	label 1	Overall
Telugu Newspapers	2792	2864	5656
English Newspapers	2136	2386	4522
Twitter (Telugu)	573	655	1228
Twitter (English)	1258	1372	2630
Twitter (Tenglish)	3538	3850	7388
Total	10297	11127	21424

Subsequently, annotators labeled the data, and finally, we aggregated the labels from all annotators into a single label.

A total of 6 journalists working for the Telugu media and are proficient in English are assigned tasks to complete the annotation, including 4 journalists of experience 3 to 5 years and two senior journalists having the experience of 10+ years. To maintain label quality and reduce subjectivity, a minimum of two annotators needed to agree for a label to be included in the dataset. In cases where the first two annotators did not agree, up to three additional annotators were assigned to annotate.

In the labeling of annotators had to choose from the four labels:

1. "True" - The provided information has significant evidence to claim as true news.

2. "Requires Advice from Senior Journalist" - The information provided requires more expertise to decide whether the information is true or false.

3. "Fake" - The provided information contradicts the fact or the information has significant evidence to claim false.

4. "Indeterminate" - There is insufficient evidence to make a clear decision on whether the information is true or false.

Data labelled as "Indeterminate" by both annotators is excluded. Text labelled as "Requires Advice from Senior Journalist" is presented to two senior journalists, who are asked to categorize the information as true, false, or indeterminate. The senior journalists independently provide labels initially, and in cases of conflicting labels, they engage in discussions to resolve difference.

We have computed inter annotator scores for the annotators Krippendorff's Alpha score as met-

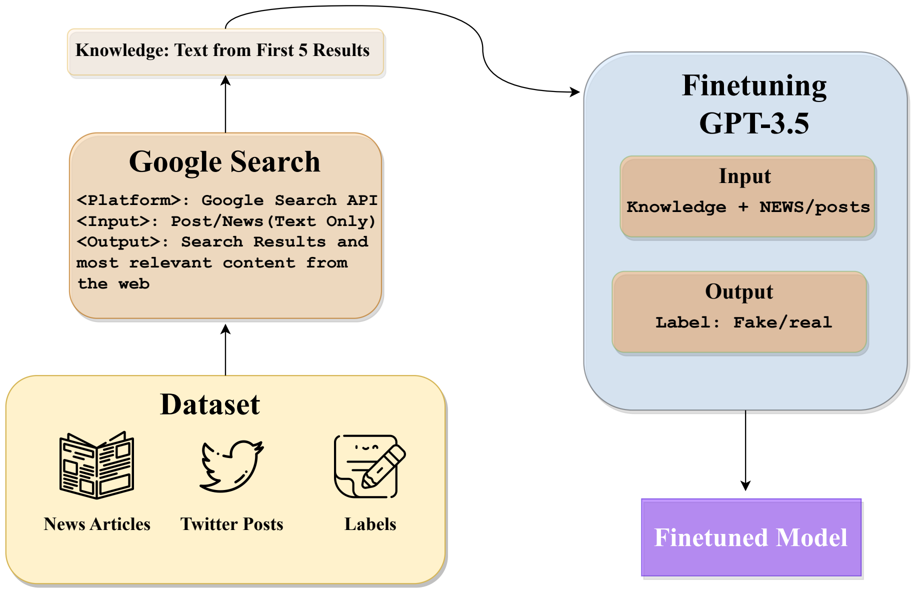

Figure 2: Finetuning of GPT-3.5 by infusing knowledge from Google

ric(Krippendorff, 2011). For the four annotators(a,b,c,d) the scores for each pair of annotators are $\alpha_{ab} = 0.863$, $\alpha_{ac} = 0.837$, $\alpha_{ad} = 0.847$, $\alpha_{bc} = 0.872$, $\alpha_{bd} = 0.861$ and $\alpha_{cd} = 0.854$, . To find out the overall agreement score, the average score for the four annotators , $\alpha = 0.856$. The inter agreement scores for the senior journalists is 0.876. The overall average score, $\alpha_T = 0.866$.

Statistics: Table I illustrates the statistics of the dataset. The dataset analysis highlights the distribution of true and false news across various sources. Telugu and English newspapers contribute a balanced representation, with both categories containing over 2000 instances each. Notably, the Tenglish Twitter category, combining Telugu and English tweets, stands out with a substantial 7388 instances, underscoring its significance as a major source of news content. This Twitter category exhibits a higher volume of both true and false news compared to traditional newspapers.

In total, the dataset comprises 21424 instances, with 10297 instances labeled as true news and 11127 instances labeled as false news. The findings underscore the necessity of source-specific considerations in addressing misinformation, as different platforms exhibit varying levels of reliability. The insights gleaned from this analysis can guide the development of more nuanced and effective strategies for detecting and mitigating misinformation in news content, particularly on dynamic platforms like Twitter.

4 Methodology

4.1 Proposed Algorithm

The proposed algorithm centers around enhancing the capabilities of GPT-3.5 to discern fake news through the integration of information gathered from Google. This strategic approach involves initiating the algorithm by forwarding the input text to Google, retrieving the top five most relevant results. These selected links serve as repositories of crucial information germane to the subject matter of the given news or text. By extracting text from these links, the algorithm gains access to insights encompassing technological advancements and societal developments. This real-time and up-to-date information proves invaluable, particularly terms unfamiliar to GPT-3.5 and tracking developments beyond its training data cut-off in 2021.

The knowledge acquired from these web results becomes an integral part of the fine-tuning process. In this phase, the text obtained from Google is seamlessly integrated with the original news input provided to GPT-3.5, as illustrated in Figure 2. The main goal during the fine-tuning is to enrich the model's understanding by incorporating the wealth of information garnered from Google. This amalgamation enhances the model's grasp of context, allowing it to better comprehend and interpret the intricacies of the information it processes. By training GPT-3.5 with insights from Google, the algorithm seeks to capitalize on the external knowledge to bolster the model's discernment capa-

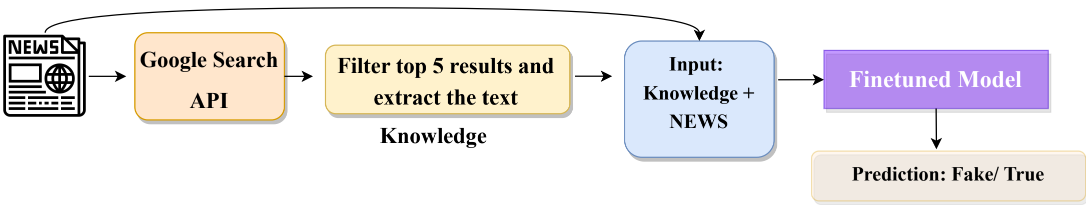

Figure 3: Flowchart depicts the usage of fine-tuned model for testing phase and in real-time applications.

bilities. This additional layer of information equips the model with a broader contextual understanding, enabling it to explore through news content more effectively and identify potential instances of misinformation. In essence, the integration of real-time information from Google serves as a dynamic enhancement strategy, addressing the evolving landscape of information beyond the initial training scope of the model.

Subsequent to the fine-tuning process, the algorithm seamlessly transitions to the analytical phase where the news or post targeted for scrutiny is dispatched to the Google API for information retrieval. This interaction initiates a process where the acquired information is systematically fed into the fine-tuned GPT-3.5 model, as visually depicted in Figure 3. Given that the model has undergone explicit training with the infusion of knowledge from Google, it manifests advanced predictive capabilities in comparison to baseline implementations.

The collaborative synergy between the fine-tuned model and the information retrieved from the Google API underscores a sophisticated approach to tackling the challenges associated with fake news detection. By leveraging external knowledge, the algorithm not only adapts to the evolving landscape of information but also enhances its analytical prowess, contributing to a more robust and effective tool for the detection of fake news and posts in the digital sphere.

To explain in detail we present the mathematical equations. η represents the news from the dataset and N_i the tokens of the news. C is the representation of the label. Assuming G is the notation for the Google search API.

$$\textit{News} : \eta = \{N_1, N_2, \ldots, N_n\} \tag{1}$$

$$c = \begin{cases} 1, & \text{if } \textit{fake news} \\ 0, & \text{if } \textit{true news} \end{cases} \tag{2}$$

$$\pi = G(N_1, N_2, \ldots, N_n) \tag{3}$$

π represents the results obtained from the Google search API with τ_i as the links. Then the text K_i is extracted from τ_i using ε function while extracts the text from the web links.

$$\pi = \{\tau_1, \tau_2, \ldots, \tau_m\} \tag{4}$$

$$K_i = \varepsilon(\tau_i) \tag{5}$$

$$K_i = \{t_1, t_2, \ldots, t_o\} \tag{6}$$

t_i are the tokens of the text K_i. From the obtained web results the text from first five results is taken and denoted as K_G. The knowledge in addition with the news is fine-tuned on GPT-3.5 with W, B as parameters of the model with loss function L_G and F represents the fine-tuned model.

$$K_G = \{K_1, K_2, \ldots, K_5\} \tag{7}$$

$$F(W, B) = \arg\min_{W,B} L_G(\{K_1 + K_2 + K_3 + K_4 + K_5 + \eta\}, c) \tag{8}$$

η_t is the news to be predicted with tokens N_{t_i}. The news is input to the Google search and the knowledge extracted is K_{G_t}.

$$\eta_t = \{N_{t_1}, N_{t_2}, \ldots, N_{t_n}\} \tag{9}$$

$$K_{G_t} = \{K_{t_1}, K_{t_2}, \ldots, K_{t_5}\} \tag{10}$$

$$P_f = F(\{K_{t_1} + K_{t_2} + K_{t_3} + K_{t_4} + K_{t_5} + \eta\}) \tag{11}$$

The knowledge and news is input to the fine-tuned model F and the models outputs the prediction P_f.

4.2 Baselines

The data is multilingual, consists of two different languages and a mixed language. So, we have opted multilingual baselines. So, that it could be suitable to evaluate. The implemented baselines are: (i) GPT 3.5 (Chen et al., 2023); (ii) GPT 3 (Brown et al., 2020);(iii) LLAMA 2(Touvron et al., 2023); (iv) multilingualBERT(Pires et al., 2019); (v)XLM-RoBERTa(Conneau et al., 2020), (vi) Integrating Google and GPT using Langchain(IGL) and (vii) few shot prompting with GPT-3.5.

For the implementation of IGL we have utilised the prompting technique the prompt is "Predict whether the following is fake news or not?: $\backslash n$ $\backslash n$ (The news/post)". For few shot prompting technique, we have prompted the GPT model by providing with five random examples from the dataset.

For the BERT-like models we have used Google Colab free GPU. For LLAMA 2 7B, 13B we have used Nvidia GPU of 108GB RAM. For the GPT models we have utilised the OpenAI API for fine-tuning and few shot prompting of the GPT-3.5.The hyperparameters used for the baselines are epoch 5, learning rate 2e-5, weight decay 0.01, frequency penalty 0, presence penalty 0.

5 Experimental Results

Table 2 presents the experimental results for the experiments performed in this study. TeluguBERT, mBERT, and XLM RoBERTa exhibited competitive performance in the detection of fake news, with accuracy values of 62.37%, 68.15%, and 69.82%, respectively. Among these, XLM RoBERTa achieved the highest accuracy. This might be because RoBERTa is multilingual and as it is enhanced form of BERT. Among the few-shot learning models, Few shot GPT-4 outperformed Few shot GPT-3.5 and IGL, achieving an accuracy of 43.57%. The latter two models demonstrated accuracy values of 41.86% and 51.49%, respectively. The IGL performed better than other prompting techniques this is because of accessing web and gains relevant up to date information.

The LLAMA models, LLAMA 2 7B and LLAMA 2 13B, exhibited superior performance compared to the previous models, achieving accuracy values of 74.15% and 75.86%, respectively. Among the GPT-3 models, GPT 3 Davinci demonstrated the highest accuracy of 75.91%, surpassing GPT 3 Babbage, GPT 3 Curie, and GPT 3 Ada, which achieved accuracy values of 74.57%, 74.36%, and 73.93%, respectively. GPT 3.5 also performed well, with an accuracy of 76.48%. As it is a LLM, pretrained on huge textual data and fine-tuned for detection of the fake news it performed better. The proposed algorithm performed much better than the baselines implemented the main reason is the algorithm learned accessing web, extracting knowledge and detecting the fake news.

Table 2: Test results: Detection of Fake News

Model	**Precision**	**Recall**	**Accuracy**
TeluguBERT	60.27	63.79	62.37
mBERT	65.75	69.56	68.15
XLM RoBERTa	66.72	70.16	69.82
Few shot GPT-3.5	40.62	43.73	41.86
Few shot GPT-4	41.40	43.82	43.57
IGL	50.13	52.32	51.49
LLAMA 2 7B	72.90	76.36	74.15
LLAMA 2 13B	73.61	77.57	75.86
GPT 3 Ada	70.67	74.20	73.93
GPT 3 Babbage	72.45	75.39	74.57
GPT 3 Curie	74.86	73.11	74.36
GPT 3 Davinci	79.26	72.43	75.91
GPT 3.5	74.51	77.06	76.48
Proposed method	79.83	82.17	**81.04**

6 Discussion

In the realm of fake news detection within the futuristic technology landscape, our proposed algorithm, leveraging the fine-tuned GPT-3.5 model with knowledge infusion from Google, outperforms other baselines that also integrate GPT-3.5 but lack dedicated fine-tuning for the specific task of fake news detection. The effectiveness of our approach is evident in its nuanced understanding of language, real-time information retrieval capabilities, and advanced contextual analysis.

One notable strength of our algorithm lies in its ability to discern speculative or sensationalized content that often eludes other baselines. For instance, when faced with a headline proclaiming "Quantum Computing Breakthrough Enables Time Travel," our algorithm excels at cross-referencing the information with recent scientific literature, expert opinions, and official announcements. The fine-tuning process ensures that it recognizes the nuances in language that may signal speculative claims, allowing it to accurately identify potential misinformation where baselines may fall short. Moreover, The fine-tuning process also equips the algorithm with a nuanced understanding of language and context, enhancing its ability to detect subtly misleading information. Consider the headline "AI Singularity Imminent: Experts Warn of Global Catastrophe." Baseline models, integrated with GPT-3.5 but lacking specific fine-tuning, may not grasp the hyperbolic nature of the claim. Our algorithm, having learned from a multitude of sources, recognizes the speculative tone and lack of substantiated evidence, contributing to a more accurate identification of this headline as potential misinformation.

Additionally, the proposed algorithm demonstrates superior performance in evaluating the credibility of news related to emerging technologies, such as blockchain or artificial intelligence. For instance, when presented with a headline asserting "Blockchain-Powered Flying Cars Set to Hit the Market Next Year", our algorithm can thoroughly analyze the feasibility of such a claim by checking for official statements from industry experts, regulatory approvals, and technological advancements. In contrast, baselines without dedicated fine-tuning for fake news detection may struggle to differentiate between credible and misleading information, relying on general language understanding without the nuanced focus our algorithm provides. Furthermore, in scenarios involving space exploration and extraterrestrial claims, our algorithm's real-time web scraping capabilities ensure that it can access the latest information from reputable sources. For example, when confronted with the headline "NASA Confirms Alien Life on Mars", our algorithm excels at cross-referencing this information with official statements and recent research findings. The dedicated fine-tuning for fake news detection enhances its ability to discern credible sources, enabling it to raise red flags when faced with sensationalized claims, a capability that might be lacking in baselines relying solely on GPT-3.5.

In the cases where knowledge retrieved from Google is incorrect: The Google is not always correct, sometimes we find blogs containing misinformation or fake news. The proposed algorithm performs much better compared to the baselines in this case. The IGL have false positives as they are context-based. As the proposed approach is fine-tuned on data, during the training phase there were data points where the knowledge from the web is incorrect/fake but where the label is true news, during these cases the web results contradicts with label, thereby creating confusion when IGL is used. As the model is fine-tuned it performed well on these cases as well.

Error Analysis:

While our proposed algorithm demonstrates notable strengths in fake news detection within the futuristic technology domain, there are few scenarios it made errors. The algorithm face challenges in distinguishing between legitimate speculation and misinformation in a rapidly evolving field. For example, if a headline speculates on the potential future capabilities of a nascent technology, such as "Experts Predict AI Will Achieve Consciousness Within a Decade," the algorithm struggles to differentiate between speculative but informed predictions and baseless claims. Balancing the understanding of speculative language while avoiding false positives poses a persistent challenge.

Another potential source of error arises when the algorithm encounters news that is related to emerging technologies with limited resources in Telugu. In this scenario the proposed algorithm showed lower performance. There are 6 predictions incorrect for every 10 posts.

From close examination of the predictions we found that the algorithm struggles to detect posts in Tenglish language. This might be because the GPT-3.5 would not have been pretrained on the Tenglish language and therefore feels difficult to understand and detect the fake news. Data augmentation or pertaining on Tenglish language would help in improving the overall performance of the model.

7 Conclusion and Future Work

In conclusion, this study addresses the issue of misinformation in the context of futuristic technology and society. Acknowledging the limitations of existing algorithms, particularly in their inability to incorporate real-time information, we proposed a novel methodology that combines the strengths of Large Language Models, specifically GPT-3.5, with the dynamic knowledge base provided by Google Knowledge. By leveraging this synergy, our framework achieved a commendable accuracy of 81.04% in detecting fake news.

The future work involves scaling the dataset to other language like Hindi, Tamil and other Indic languages. We would also like to develop a pretraining model especially for Tenglish as it is expected to perform better. We would like to develop a dataset in Telugu that also involves the fake news on investments.

Limitations

This approach is specifically designed for handling textual data, ensuring optimized performance for text-based processing. Since it focuses exclusively on text, image data is not included in the dataset, allowing for a more streamlined and efficient analysis. To leverage advanced AI capabilities, we utilize Google API and OpenAI models, which operate on a structured billing model. This aligns with standard industry practices for accessing state-of-the-art machine learning services. While these models are closed-source, they provide reliable and high-quality performance for text processing.

Ethics Statement

Our primary goal is to detect fake news while ensuring that the reputation of sources remains unaffected. To maintain anonymity, we have rephrased the collected data, preventing any potential reputational impact on sources or users. Additionally, we strongly oppose any misuse of the dataset for generating or spreading fake news.

References

Usman Ahmed, Jerry Chun-Wei Lin*, and Vicente Garcia Diaz. 2022. Automatically temporal labeled data generation using positional lexicon expansion for focus time estimation of news articles. *ACM Transactions on Asian and Low-Resource Language Information Processing*.

Arshia Arya, Saloni Dash, Syeda Zainab Akbar, Joyojeet Pal, and Anirban Sen. 2022. Poster: Leveraging question answering to understand context specific patterns in fact checked articles in the global south. In *ACM SIGCAS/SIGCHI Conference on Computing and Sustainable Societies (COMPASS)*, pages 628–631.

J Briskilal, Ch VM Sai Praneeth, Ch Chaitanya, M Jaya Karthik, and P Purnachandra Reddy. 2023. An ensemble method to classify telugu idiomatic sentences using deep learning models. In *2023 International Conference on Inventive Computation Technologies (ICICT)*, pages 65–71. IEEE.

Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. Language models are few-shot learners.

Hangjun Che, Baicheng Pan, Man-Fai Leung, Yuting Cao, and Zheng Yan. 2023. Tensor factorization with sparse and graph regularization for fake news detection on social networks. *IEEE Transactions on Computational Social Systems*.

Xuanting Chen, Junjie Ye, Can Zu, Nuo Xu, Rui Zheng, Minlong Peng, Jie Zhou, Tao Gui, Qi Zhang, and Xuanjing Huang. 2023. How robust is gpt-3.5 to predecessors? a comprehensive study on language understanding tasks. *arXiv preprint arXiv:2303.00293*.

Thomas Clark, Costanza Conforti, Fangyu Liu, Zaiqiao Meng, Ehsan Shareghi, and Nigel Collier. 2021. Integrating transformers and knowledge graphs for twitter stance detection. In *Proceedings of the Seventh Workshop on Noisy User-generated Text (W-NUT 2021)*, pages 304–312.

Alexis Conneau, Kartikay Khandelwal, Naman Goyal, Vishrav Chaudhary, Guillaume Wenzek, Francisco Guzmán, Edouard Grave, Myle Ott, Luke Zettlemoyer, and Veselin Stoyanov. 2020. Unsupervised cross-lingual representation learning at scale. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8440–8451, Online. Association for Computational Linguistics.

Yasan Ding, Bin Guo, Yan Liu, Yunji Liang, Haocheng Shen, and Zhiwen Yu. 2022. Metadetector: Meta event knowledge transfer for fake news detection. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 13(6):1–25.

Huong T Duong, Van H Ho, and Phuc Do. 2023. Fact-checking vietnamese information using knowledge graph, datalog, and kg-bert. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(10):1–23.

Qian Han, Cristian Molinaro, Antonio Picariello, Giancarlo Sperli, Venkatramanan S Subrahmanian, and Yanhai Xiong. 2021. Generating fake documents using probabilistic logic graphs. *IEEE Transactions on Dependable and Secure Computing*, 19(4):2428–2441.

Klaus Krippendorff. 2011. Computing krippendorff's alpha-reliability.

Claudio Marche, Ilaria Cabiddu, Christian Giovanni Castangia, Luigi Serreli, and Michele Nitti. 2023. Implementation of a multi-approach fake news detector and of a trust management model for news sources. *IEEE Transactions on Services Computing*.

Telmo Pires, Eva Schlinger, and Dan Garrette. 2019. How multilingual is multilingual BERT? In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4996–5001, Florence, Italy. Association for Computational Linguistics.

Eduri Raja, Badal Soni, and Samir Kumar Borgohain. 2023. Fake news detection in dravidian languages using transfer learning with adaptive finetuning. *Engineering Applications of Artificial Intelligence*, 126:106877.

Simone Raponi, Zeinab Khalifa, Gabriele Oligeri, and Roberto Di Pietro. 2022. Fake news propagation: a review of epidemic models, datasets, and insights. *ACM Transactions on the Web (TWEB)*, 16(3):1–34.

Georg Rehm, Julian Moreno-Schneider, and Peter Bourgonje. 2018. Automatic and manual web annotations in an infrastructure to handle fake news and other online media phenomena. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).

Qingyun Ren, Bingyin Zhou, Dongli Yan, and Wei Guo. 2023. Fake news classification using tensor decomposition and graph convolutional network. *IEEE Transactions on Computational Social Systems*.

Mina Schütz. 2023. Disinformation detection: Knowledge infusion with transfer learning and visualizations. In *European Conference on Information Retrieval*, pages 468–475. Springer.

Shivangi Singhal, Rajiv Ratn Shah, and Ponnurangam Kumaraguru. 2022. Factdrill: A data repository of fact-checked social media content to study fake news incidents in india. In *Proceedings of the international AAAI conference on web and social media*, volume 16, pages 1322–1331.

Chetana B Thaokar, Mayur Rathod, Shayeek Ahmed, Jitendra Kumar Rout, and Minakhi Rout. 2022. A multi-linguistic fake news detector on hindi, marathi and telugu. In *2022 OITS International Conference on Information Technology (OCIT)*, pages 324–329. IEEE.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. 2023. Llama: Open and efficient foundation language models. *arXiv preprint arXiv:2302.13971*.

Jinxia Wang, Stanislav Makowski, Alan Cieślik, Haibin Lv, and Zhihan Lv. 2023. Fake news in virtual community, virtual society, and metaverse: A survey. *IEEE Transactions on Computational Social Systems*.

Bingbing Xie, Xiaoxiao Ma, Jia Wu, Jian Yang, and Hao Fan. 2023. Knowledge graph enhanced heterogeneous graph neural network for fake news detection. *IEEE Transactions on Consumer Electronics*.

Xingsi Xue and Wenyu Liu. 2023. Integrating heterogeneous ontologies in asian languages through compact genetic algorithm with annealing re-sample inheritance mechanism. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 22(3):1–21.

Mesay Gemeda Yigezu, Selam Kanta, Olga Kolesnikova, Grigori Sidorov, and Alexander Gelbukh. 2023. Habesha@ dravidianlangtech: Abusive comment detection using deep learning approach. In *Proceedings of the Third Workshop on Speech and Language Technologies for Dravidian Languages*, pages 244–249.

Biru Zhu, Xingyao Zhang, Ming Gu, and Yangdong Deng. 2021. Knowledge enhanced fact checking and verification. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 29:3132–3143.

Yongchun Zhu, Qiang Sheng, Juan Cao, Qiong Nan, Kai Shu, Minghui Wu, Jindong Wang, and Fuzhen Zhuang. 2022. Memory-guided multi-view multi-domain fake news detection. *IEEE Transactions on Knowledge and Data Engineering*.