Attention-Based Deep Learning Models for Detecting Misinformation of Long-Term Effects of COVID-19

Jian-An Chen
Institute of Biomedical Informatics
National Yang Ming
Chiao Tung University
Taipei, Taiwan
masachen.md11@nycu.edu.tw

Che-Lun Hung
Institute of Biomedical Informatics
National Yang Ming
Chiao Tung University
Taipei, Taiwan
clhung@nycu.edu.tw

Chun-Ying Wu
Institute of Biomedical Informatics
National Yang Ming
Chiao Tung University
Taipei, Taiwan
cywu4@nycu.edu.tw

***Abstract*—During the COVID-19 pandemic, the surge of misinformation on social media threatens public understanding and epidemic prevention policies. Even as the pandemic is being controlled, long-term COVID-19 and reinfection risks still need to be included in COVID-19 policies and information. This study presented a deep learning approach to detect fake news related to the long-term influences of COVID-19. The data is collected and refined from reliable open sources with data processing techniques. Then, the various attention-based deep learning models like HAN, BERT, and XLNet are trained to detect misinformation about the long-term effects of COVID-19 based on the collected data. The F1 score reached 94.96%, showing the strong performance of the deep learning models. The method demonstrated high effectiveness in identifying such false content, contributing automatic tools for detecting misinformation on the long-term impacts of the COVID-19 pandemic.**

***Index Terms*—Attention-based models, Misinformation, COVID-19, Pre-trained language models (PLMs)**

I. Introduction

From 2019 to 2022, the world experienced the coronavirus disease 2019 (COVID-19) pandemic. Governments worldwide and the World Health Organization (WHO) worked hard to control the spread of the virus. Amidst this global health crisis, the demand for reliable information sources and accurate health advice has increased. However, with the growing needs, misinformation and false news spread rapidly on social media, causing public confusion. The WHO described the spread of misinformation as an “infodemic” [1]. They pointed out that such misinformation could threaten national epidemic prevention policies. When the public trusts incorrect or misleading information, it can lead to adverse health behaviors and non-compliance with health policies, further exacerbating the pandemic. With the widespread distribution of COVID-19 vaccines, the pandemic was gradually controlled. However, COVID-19 did not disappear. Instead, there emerged post-infection symptoms known as long COVID, which have been confirmed in at least 10% of those who contracted the virus [2]. Additionally, some patients were reinfected with COVID-19 after their initial recovery, a phenomenon termed ”reinfection”. Based on research analyzing data from the U.S. Department of Veterans Affairs, reinfections have been shown to increase the risk of mortality, hospitalization, and post-symptomatic conditions for patients [3].

Though the immediate threat of COVID-19 is fading, the potential risks brought by long COVID and reinfection require the public to continuously focus on COVID-19-related policies and information. This sets the stage for a post-pandemic era co-existing with the virus. Even as the pandemic is being controlled and countries transition into the post-pandemic era, the issues of fake news and misinformation have not disappeared. Specifically concerning long COVID and reinfection, such topics remain a focal point for misinformation. Thus, the practical and swift identification and classification of such fake news becomes crucial. With the advancements in Natural Language Processing (NLP) and deep learning technologies, this study explores the performance of various deep learning models in identifying fake news. The aim is to offer a more scientific and efficient method for fake news detection in the post-pandemic era. Texts related to long COVID and reinfection were gathered from open-source databases and through web crawling. The gathered data underwent a preprocessing phase to clean and refine it. Various machine learning and deep learning models were trained and evaluated after preprocessing based on their performance. Finally, we analyzed the classification efficiency of deep learning models in distinguishing fake news. The F1 score of 94.96% emphasizes the robust performance of the attention-based models. The results exhibit the effectiveness of attention-based models in distinguishing misinformation.

II. Related work

During the COVID-19 pandemic, many research studies have used machine learning and deep learning techniques to detect fake news and address the issue of misinformation.

Patwa et al. [4] gathered COVID-19-related texts from publicly available fact-checking websites and social media

979-8-3503-5409-6/24/$31.00 ©2024 IEEE
DOI 10.1109/CAI59869.2024.00053

platforms. They used TF-IDF for feature extraction and different machine learning algorithms, including logistic regression, support vector machines, decision trees, and gradient boosting, for the binary classification of fake news. Das et al. [5] utilized pre-trained language models (PLMs) such as RoBERTa and XLNet for preprocessing and training on the same dataset. Ensembling predictions from multiple models through voting, they achieved admirable results in the CONSTRAINT2021 COVID-19 Fake News Detection competition.

Paka et al. [6] believed that more than relying on textual features might be required for accurate fake news classification. Therefore, they collected tweets related to COVID-19 from Twitter. Beyond the text content, they also gathered data such as the number of likes for a tweet, URL links, and the follower count of the tweeters. Using a cross-stitch unit combined with an LSTM architecture, they proposed a multi-feature classification approach for fake news.

Furthermore, teams of Chinese speakers have also devoted themselves to this research domain. They employed deep learning frameworks such as RNN, CNN, and Transformer to classify COVID-19 fake news in Chinese text [7]. These efforts underscore the global emphasis on the infodemic, striving to ensure the public receives accurate and reliable information.

In addition, large language models(LLMs) based on transformers such as ChatGPT [8] have blossomed in recent years. These models can understand natural language and interact with users, which may contribute to developing more robust tools for combating the spread of false news.

III. Methods

The methods consist of four primary steps: data collection, data preprocessing, data analysis, and modeling. First, data was gathered from various publicly accessible online sources. Due to the inconsistency of online sources, this data underwent a preprocessing phase to clean and refine it. After preprocessing, a fundamental analysis was conducted to understand the dataset's characteristics further. Based on these insights, various machine learning and deep learning models were trained and evaluated according to their performance.

A. Data collection

We collected articles and claims related to COVID-19 from various online sources. These materials were then filtered using keywords associated with long COVID and reinfection. The keywords included chronic, long COVID, long COVID, longcovid, long-term, persistent, after-effects, sequelae, complications, recovery, post covid, post-covid, omicron, subvariant, reinfection, immune, and variant. The resulting collection formed the dataset for our experiments. Each piece of data was categorized as either "genuine" or "fake" and gathered from three primary sources:

Open Source Dataset

Fighting an Infodemic [4]: An earlier dataset includes topics related to COVID-19 sourced from platforms like Twitter, Facebook, and fact-checking websites. This dataset was chosen for the *Constraint@AAAI2021 - COVID19 Fake News Detection in English competition*. Only labeled data from this dataset were used, available on GitHub [9].

CTF (COVID-19 Twitter Fake News) [6]: CTF is a dataset that focuses on tweets from the large social media platform Twitter. It contains both labeled and unlabeled data concerning genuine and fake COVID-19 news. In addition to the content, this dataset includes the tweeters' user-profile features. For this study, only the labeled text content data was used.

CoAID (Covid-19 heAlthcare mIsinformation Dataset) [10]: CoAID is a diverse COVID-19 fake news dataset containing news from the internet and social media platforms, user engagements related to these fake messages, and tweets and labels appearing on Twitter.

FibVID (Fake news information-broadcasting dataset of COVID-19) [11]: This dataset collects claims from fact-checking websites like Snopes and Politifact and related discourse from Twitter. It also gathers user profile information from social media platforms. This dataset includes COVID and non-COVID topics, divided into four labels. Only data related to COVID-19 from the categories 0 and 1 was used.

FaCOV [12]: Collected from 13 English fact-checking websites related to COVID-19, this dataset includes article titles, URLs, claims made within the articles, and abstracts. The titles and article contents were merged, and data with two category labels were utilized in the experiments.

FactCheck Website

While open-source databases offer significant help, they are often limited to data from 2019 and 2021. To overcome this limitation, our study employed web scraping and data cleaning techniques to gather more recent data, facilitating deeper analysis and investigation. The data collection process comprises the following steps:

Source: Snopes [13] and PolitiFact [14] have been verified by the International Fact-Checking Network (IFCN). This certification indicates that the data and labels from these sites underwent transparent and rigorous validation processes, enhancing data reliability. Therefore, both of them were selected as primary data sources.

Extraction: Considering the extensive amount of target data, we used web scraping to systematically extract articles tagged as "CORONAVIRUS" and "COVID-19" from Snopes (data up to August 31, 2023) and PolitiFact (data up to July 31, 2023). Alongside the contents, labels were also collected for model training. From Snopes and PolitiFact, 1500 and 806 texts were extracted, respectively. After that, collected data were filtered by keywords to align more closely with the topic.

Labeling: In PolitiFact, articles are categorized into six labels: pants-on-fire, false, barely-true, half-true, mostly-true, and true. Snopes, on the other hand, classifies articles into fourteen labels: true, mostly-true, mixture, mostly-false, false, unproven, outdated, miscaptioned, correct-attribution, misattributed, scam, legend, labeled-satire, and lost-legend. According to the research by Khan et al. [15], the labels from different sources were reclassified into two categories:

TABLE I
SAMPLE SIZE FROM DIFFERENT SOURCES

Source	Time until	Samples	Fake	Genuine
CTF	~2021	1292	1130	162
Fighting an Infodemic	~2021	218	62	156
CoAID	~2020	70	0	70
FibVID	~2020	615	318	297
FaCOV	~2021	811	811	0
PolitiFact	~2023	87	42	45
Snopes	~2023	15	9	6
CDC+WHO	~2023	58	0	58
Total		3166	2372	794

genuine and fake.

Governmental Bodies

Highly reputable government and public institution websites, such as the World Health Organization (WHO) and the Centers for Disease Control and Prevention (CDC), were also considered primary sources. These institutions have consistently provided up-to-date information and guidelines throughout the pandemic, and as a result, they are widely considered reliable and accurate data sources. We gathered the articles related to "long COVID" and "Reinfection" from the COVID-19 sections of these websites. Given that the original textual data might be excessively lengthy or contain superfluous details, we used ChatGPT [8] to organize and refine the acquired content. Using ChatGPT, we structured the texts to ensure each claim had appropriate length and clarity. Subsequently, each claim was labeled as "genuine" due to its reputable source. The dataset used for model training contained the latest information through these steps.

So far, the filtered sample counts from various data sources are presented in Table I.

B. Data preprocessing

When merging multiple open-source datasets, ensuring that the assembled dataset does not contain duplicate entries is vital. We cross-referenced the data with previously known open-source datasets. All duplicate samples identified were removed to prevent redundancy, which could influence the performance of subsequent model training and analysis.

Social media posts and articles often contain emojis and external links (URLs) that do not add significant value in distinguishing between genuine and fake news. Therefore, the Tweet-preprocessor package was utilized during the preprocessing phase to remove emojis and URLs from the texts.

The labels "genuine" and "fake" were encoded as "0" and "1" respectively. Fig. 1 shows the distribution of our dataset. Due to the public data sources being skewed towards the "fake" category, the dataset displayed an imbalanced label distribution. Thus, we adopted stratified sampling, using 10% of the data for testing and the remaining 90% for training. The sampling method ensures that the label proportions in both the test and training sets remain consistent, preventing cases in which random sampling might lead to an insufficient number of "genuine" samples.

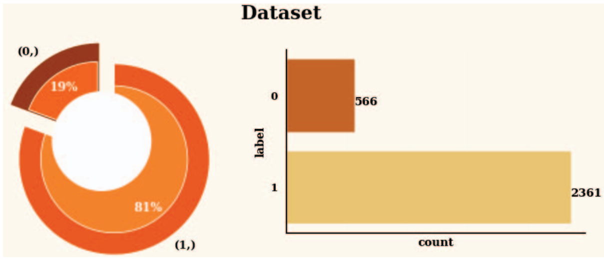

Fig. 1. Label distribution after preprocessing

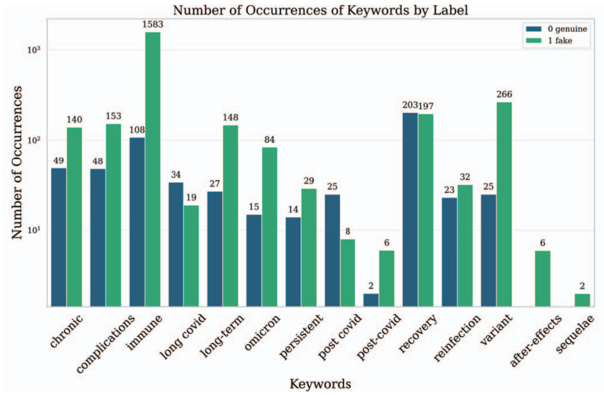

Fig. 2. Number of keywords occurrences by label

C. Data analysis

1) Keywords occurrences: By analyzing the frequency of specific keywords in the data, we can understand the public's focus. In Fig. 2, more than half of the fake samples contain the term "immune," but only a few are in the genuine category. The term "recovery" is primarily seen in genuine samples; however, a similar proportion of fake samples also use it. Other keywords such as "variant," "complication," and "chronic" also appear at higher frequencies in fake samples. This suggests that such terms are often used to propagate fake news.

2) Sentiment analysis: Sentiment analysis can indicate whether the content of an article is positive, negative, or neutral. Moreover, it can assess the subjectivity of the text, determining whether the information is based on facts or the author's personal opinions. In order to get a deeper understanding of the textual data collected from open-source databases and fact-checking websites, we utilized the TextBlob package [16] for sentiment analysis. TextBlob's sentiment analysis provides a polarity value ranging from -1 to 1, signifying the sentiment from entirely negative to entirely positive. To differentiate between various sentiments, based on polarity, we categorized sentiment into the following five groups:

- Strongly Negative: polarity values between -1 and -0.5.
- Slightly Negative: polarity values between -0.5 and -0.1.
- Neutral: polarity values between -0.1 and 0.1.
- Slightly Positive: polarity values between 0.1 and 0.5.
- Strongly Positive: polarity values between 0.5 and 1.

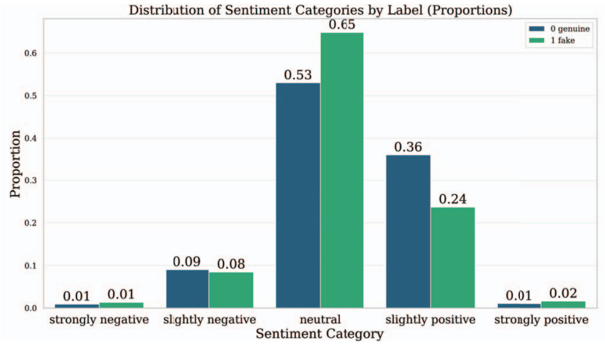

Fig. 3. Data distribution (percentage) of different sentiment polarity groups

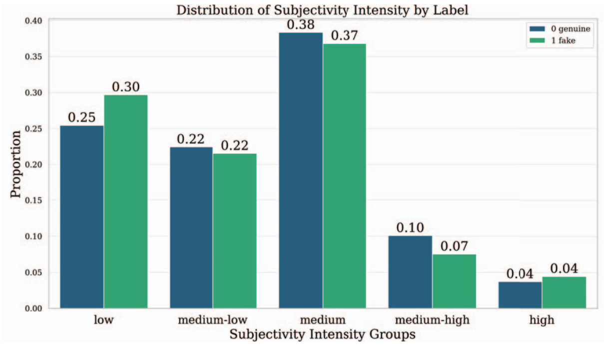

Fig. 4. Data distribution (percentage) of different subjectivity groups

As illustrated in Fig. 3, whether the text is fake or genuine, most content is primarily categorized as "neutral." 65% of the content in fake texts falls into this category, while 53% of the content in genuine texts does so. Interestingly, in the "slightly positive" category, genuine texts (accounting for 36%) surpass the 24% in fake texts, suggesting that genuine content frequently uses favorable terms. As for the other sentiment categories ("strongly negative," "slightly negative," and "strongly positive"), the distribution between the two labels is similar, with no significant differences.

Apart from sentiment analysis, TextBlob also provides a subjectivity value ranging from 0 to 1, representing the spectrum from entirely objective to entirely subjective. We categorized the degree of subjectivity based on the value into the following five groups to make it easier to understand:

- Low Subjectivity: values between 0 and 0.2.
- Medium-Low Subjectivity: values between 0.2 and 0.4.
- Medium Subjectivity: values between 0.4 and 0.6.
- Medium-High Subjectivity: values between 0.6 and 0.8.
- High Subjectivity: values between 0.8 and 1.

From the analysis presented in Fig. 4, the distribution between genuine and fake texts appears very similar, primarily concentrated in the "medium subjectivity" category. Genuine texts have 38% of their content in this category, while fake texts contain 37%. Notably, the "low subjectivity" category has 5% more fake texts than genuine ones. Conversely, in the "medium-high subjectivity" category, genuine texts exceed fake texts by a margin of 3%.

Considering the analysis of sentiment and subjectivity, even though there are minor differences in the distribution between genuine and fake texts, these differences may not be sufficient to serve as clear classification criteria. Moreover, sentiment analysis still faces challenges in natural language, such as accurately detecting sarcasm. As a result, more precise methodologies, like machine learning algorithms and deep learning models, are needed to aid us in distinguishing genuine from fake news concerning long COVID and reinfections.

D. Models

SVM

This study first chose SVM (Support Vector Machine) [17] as the baseline model. In text classification, linear classifiers are often regarded as excellent baselines. Comparing the linear classifier with deep models can confirm whether they are correctly fine-tuned and employed [18]. To employ SVM for classification tasks, we generated uni-gram TF-IDF features from our training set data and trained the SVM using these features for binary classification.

Attention-based models

BERT [19] BERT (Bidirectional Encoder Representations from Transformers) is a pre-trained language model based on the Transformer architecture introduced by Google in 2018. It is trained on vast amounts of text using masked language modeling and next-sentence prediction. Through self-attention mechanisms and contextual information, BERT produces contextualized word representations. It is one of the most famous language models widely employed across various NLP tasks.

RoBERTa [20] Based on BERT, RoBERTa (A Robustly Optimized BERT Pretraining Approach) was proposed by Facebook in 2019. It was trained on more data. Unlike BERT, RoBERTa removed the next sentence prediction pre-training task and integrated optimization strategies to enhance the original BERT model, such as using larger batch sizes and dynamic masking. The original study verified that RoBERTa outperformed BERT in various NLP tasks.

DeBERTa [21] DeBERTa (Decoding-enhanced BERT with Disentangled Attention) introduces a novel attention mechanism known as "disentangled attention" to refine the original self-attention. It also employs an enhanced mask decoder and integrates token absolute positioning, which aids the model in capturing sequence information more effectively and improving the pre-training process. Across multiple NLP tasks, DeBERTa has demonstrated superior performance, outshining other pre-trained models, including BERT and RoBERTa.

HAN [22] HAN (Hierarchical Attention Networks) incorporates attention mechanisms at multiple levels, specifically at the word and sentence levels, to capture diverse hierarchical structures of documents. The authors employed RNNs combined with word-attention and sentence-attention layers for text classification. This architecture yielded state-of-the-art results across six datasets at that time.

XLNet [23] XLNet is a pre-trained language model presented by Google in 2019. XLNet proposed permutation language modeling to achieve bidirectional contextual comprehension. With the Transformer-XL architecture adept at training on significant texts, XLNet conquered the challenges of long-text understanding. It refines some limitations present during BERT's pre-training, demonstrating its superior performance against models like BERT and RoBERTa across multiple NLP benchmark tasks.

GPT-4 [24] GPT-4 is a Transformer-based auto-regressive model pre-trained to predict the next token. It demonstrates human-level performance on various professional and academic benchmarks, including achieving a score around the top 10% in a simulated bar exam. With Open AI API, GPT-4 can be accessed to make predictions and output the answer in JSON format.

IV. Experiments and Results

We utilized Scikit-learn [25] to employ 5-fold cross-validation for training the SVM and selected the best model for final evaluation. To fine-tune PLMs, we utilized checkpoints provided by Hugging Face. PLMs were trained using the AdamW optimizer and cross-entropy loss, setting the learning rate to 2e-5 and training 20 epochs. HAN was trained for 20 epochs with its original settings. We executed the same training procedure five times using five different random seeds and selected models for final evaluation based on the highest validation F1-score [26]. HAN was trained on the Tesla T4; all other PLMs were trained on the RTX A5000.

Model performances were evaluated on famous metrics, including accuracy, precision, recall, F1-score, and AUC(Area Under the ROC Curve). In Table II, the SVM exhibited an accuracy of 89.08%, a precision of 91.46%, recall of 95.34%, F1-score of 93.36%, and an AUC of 92.02%. The BERT model outperformed SVM with an accuracy of 91.13%, precision of 93.75%, and an impressive AUC of 96.31%, sharing the same recall of 95.34% with an F1-score of 94.54%. RoBERTa, which showed an accuracy of 91.81%, achieved the highest AUC among the models at 96.56%. Similarly, DeBERTa matched RoBERTa's accuracy and mirrored its recall of 95.76% but slightly trailed in AUC with 95.75%. HAN presented an accuracy of 90.78% and an AUC of 93.09%, while XLNet aligned with RoBERTa and DeBERTa in accuracy at 91.81% and featured an AUC of 96.01%. GPT-4 showed an accuracy of 82.25%, precision of 91.82%, recall of 85.59%, and F1-score of 88.60%. The AUC is not feasible from GPT-4 results, as AUC requires the model to predict the probability distribution for different labels, which is not directly provided by GPT-4.

Attention-based models, particularly PLMs, demonstrated superior performance. XLNet and DeBERTa stood out with identical classification results, achieving the highest accuracy, recall, and F1-score. However, XLNet outperformed DeBERTa in AUC. While RoBERTa had slightly lower recall and F1-score than XLNet, it excelled in precision and AUC.

TABLE II
Accuracy, Precision, Recall, F1-score and AUC of models on the test data.

Model	Accuracy	Precision	Recall	F1-score	AUC
SVM	89.08%	91.46%	95.34%	93.36%	92.02%
BERT	91.13%	93.75%	95.34%	94.54%	96.31%
RoBERTa	**91.81%**	**94.54%**	95.34%	94.94%	**96.56%**
DeBERTa	**91.81%**	94.17%	**95.76%**	**94.96%**	95.75%
HAN	90.78%	94.09%	94.49%	94.29%	93.09%
XLNet	**91.81%**	94.17%	**95.76%**	**94.96%**	96.01%
GPT-4	82.25%	91.82%	85.59%	88.60%	N/A

Fig. 5 shows the confusion matrix of final predictions from SVM and XLNet on the test set. Since the models were trained on an imbalanced dataset, the Linear method showed some tendency to predict "1". However, interestingly, XLNet did not exhibit this type of tendency, suggesting that it may have inherent robustness against the challenges from imbalanced datasets. Table III is the actual case of using XLNet to detect fake and genuine news that was not included in the test set.

V. Conclusions & Discussion

This study utilized open-source databases and reputable websites to gather text data about long COVID and reinfections. We could understand the characteristics of articles and claims surrounding these topics through information engineering techniques. AI models, particularly attention-based models and linear classifiers, have proven they can effectively detect misleading or inaccurate information. Thus, such AI models can serve as a tool to help the public differentiate between genuine and fake information. Incorporating the GPT-4 model into the comparison, using GPT-4 for predictions directly did not work as well as the other models. It demonstrated the need to give GPT-4 more training to make it more suitable for specific tasks. This process should help fix the current issues and make GPT-4 a more practical tool.

Furthermore, the experiment results demonstrate that even training solely on textual content can achieve high prediction accuracy. However, it is still important to avoid misclassifications. Ensemble methods that combine predictions from multiple models may help mitigate the risk of misclassifications by fusing different models' perspectives and strengths. In addition, a challenge encountered during experiments was the data imbalance. It might suggest that fake information is spreading more widely on the internet. Gathering more genuine and recent data is still necessary to address this issue.

Acknowledgment

This study is financially supported by the National Science and Technology Council (NSTC) of Taiwan under Grants NSTC 112-2314-B-A49-049-MY3, 112-2634-F-A49-003, and 113-2321-B-A49-011.

References

[1] "Infodemic." Accessed: Sep. 28, 2023. [Online]. Available: https://www.who.int/health-topics/infodemic

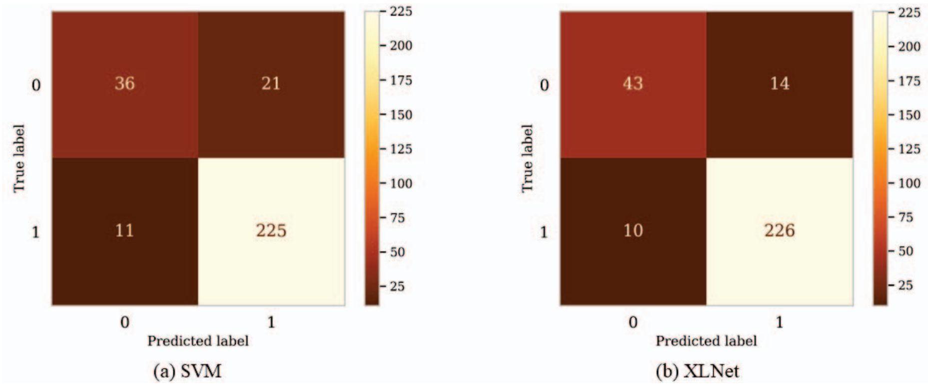

Fig. 5. Confusion matrix of (a) SVM model and (b) XLNet on the test data.

TABLE III
REAL CASE INFERENCE USING XLNET.

Content	Date	Model prediction	Ground truth
A German study has revealed long COVID is linked to the vaccine.	November 2, 2023	fake	fake
Long Covid is just a side effect of Covid-19 vaccinations.	November 29, 2022	fake	fake
Covid vaccination before infection strongly linked to reduced risk of developing long covid	November 23, 2023	genuine	genuine

[2] H. E. Davis, L. McCorkell, J. M. Vogel, and E. J. Topol, "Long COVID: major findings, mechanisms and recommendations," Nat. Rev. Microbiol., vol. 21, no. 3, Art. no. 3, Mar. 2023, doi: 10.1038/s41579-022-00846-2.

[3] B. Bowe, Y. Xie, and Z. Al-Aly, "Acute and postacute sequelae associated with SARS-CoV-2 reinfection," Nat. Med., vol. 28, no. 11, Art. no. 11, Nov. 2022, doi: 10.1038/s41591-022-02051-3.

[4] P. Patwa et al., "Fighting an Infodemic: COVID-19 Fake News Dataset," vol. 1402, 2021, pp. 21–29. doi: 10.1007/978-3-030-73696-5_3.

[5] S. D. Das, A. Basak, and S. Dutta, "A Heuristic-driven Ensemble Framework for COVID-19 Fake News Detection." arXiv, Jan. 10, 2021. doi: 10.48550/arXiv.2101.03545.

[6] W. S. Paka, R. Bansal, A. Kaushik, S. Sengupta, and T. Chakraborty, "Cross-SEAN: A cross-stitch semi-supervised neural attention model for COVID-19 fake news detection," Appl. Soft Comput., vol. 107, p. 107393, Aug. 2021, doi: 10.1016/j.asoc.2021.107393.

[7] C. Yang, X. Zhou, and R. Zafarani, "CHECKED: Chinese COVID-19 fake news dataset," Soc. Netw. Anal. Min., vol. 11, no. 1, p. 58, Jun. 2021, doi: 10.1007/s13278-021-00766-8.

[8] "ChatGPT." Accessed: Oct. 19, 2023. [Online]. Available: https://chat.openai.com

[9] "covid_fake_news/data at main · diptamath/covid_fake_news," GitHub. Accessed: Oct. 18, 2023. [Online]. Available: https://github.com/diptamath/covid_fake_news/tree/main/data

[10] L. Cui and D. Lee, "CoAID: COVID-19 Healthcare Misinformation Dataset." arXiv, Nov. 03, 2020. doi: 10.48550/arXiv.2006.00885.

[11] J. Kim, J. Aum, S. Lee, Y. Jang, E. Park, and D. Choi, "FibVID: Comprehensive fake news diffusion dataset during the COVID-19 period," Telemat. Inform., vol. 64, p. 101688, Nov. 2021, doi: 10.1016/j.tele.2021.101688.

[12] S. Sharma, E. Agrawal, R. Sharma, and A. Datta, "FaCov: COVID-19 Viral News and Rumors Fact-Check Articles Dataset," Proc. Int. AAAI Conf. Web Soc. Media, vol. 16, pp. 1312–1321, May 2022, doi: 10.1609/icwsm.v16i1.19383.

[13] "COVID Archives — Snopes.com." Accessed: Oct. 19, 2023. [Online]. Available: https://www.snopes.com/tag/covid-19/

[14] "Fact-checks — PolitiFact." Accessed: Oct. 19, 2023. [Online]. Available: https://www.politifact.com/factchecks/list/?category=coronavirus

[15] J. Y. Khan, Md. T. I. Khondaker, S. Afroz, G. Uddin, and A. Iqbal, "A benchmark study of machine learning models for online fake news detection," Mach. Learn. Appl., vol. 4, p. 100032, Jun. 2021, doi: 10.1016/j.mlwa.2021.100032.

[16] "TextBlob: Simplified Text Processing — TextBlob 0.16.0 documentation." Accessed: Oct. 12, 2023. [Online]. Available: https://textblob.readthedocs.io/en/dev/

[17] C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.

[18] Y.-C. Lin, S.-A. Chen, J.-J. Liu, and C.-J. Lin, "Linear Classifier: An Often-Forgotten Baseline for Text Classification," in Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 1876–1888. doi: 10.18653/v1/2023.acl-short.160.

[19] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." arXiv, May 24, 2019. doi: 10.48550/arXiv.1810.04805.

[20] Y. Liu et al., "RoBERTa: A Robustly Optimized BERT Pretraining Approach." arXiv, Jul. 26, 2019. doi: 10.48550/arXiv.1907.11692.

[21] P. He, X. Liu, J. Gao, and W. Chen, "DeBERTa: Decoding-enhanced BERT with Disentangled Attention." arXiv, Oct. 06, 2021. doi: 10.48550/arXiv.2006.03654.

[22] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical Attention Networks for Document Classification," in Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, San Diego, California: Association for Computational Linguistics, 2016, pp. 1480–1489. doi: 10.18653/v1/N16-1174.

[23] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le, "XLNet: Generalized Autoregressive Pretraining for Language Understanding," in Advances in Neural Information Processing Systems, Curran Associates, Inc., 2019. Accessed: Jul. 12, 2023. [Online]. Available: https://papers.nips.cc/paper_files/paper/2019/hash/dc6a7e655d7e5840e66733e9ee67cc69-Abstract.html

[24] OpenAI, "GPT-4 Technical Report." arXiv, Mar. 27, 2023. Accessed: Dec. 04, 2023. [Online]. Available: http://arxiv.org/abs/2303.08774

[25] F. Pedregosa et al., "Scikit-learn: Machine Learning in Python," J. Mach. Learn. Res., vol. 12, no. 85, pp. 2825–2830, 2011.

[26] I. Chalkidis et al., "LexGLUE: A Benchmark Dataset for Legal Language Understanding in English," in Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Dublin, Ireland: Association for Computational Linguistics, May 2022, pp. 4310–4330. doi: 10.18653/v1/2022.acl-long.297.