Journal of Computer Science

Research Paper

Detecting Disinformation: Enriching the COVAX Reality Dataset with BERT for Stance Detection

Saad Ahmed[1], Asma Qaiser[1], Sidrah Abdullah[2], and Muhammad Shoaib Siddiqui[3]

[1]*Department of Computer Science, FEST, IQRA University Karachi, Pakistan*
[2]*Computer Science, Information Technology Department, NED University of Engineering and Technology, Karachi, Pakistan*
[3]*Faculty of Computer and Information Systems, Islamic University of Madinah, Madinah, Saudi Arabia*

Article history
Received: 16 January 2025
Revised: 15 April 2025
Accepted: 20 May 2025

*Corresponding Author:
Saad Ahmed
Department of Computer Science, FEST, IQRA University Karachi, Pakistan
Email: support@thescipub.com

Abstract: Social media news pages and online news portals have significantly grown in popularity among the users across the globe for accessing the latest and important news. Unlike traditional newspapers, online news portals and social media provide the news round the clock and can be accessed from any location free of cost. However, the internet sources, especially the social media, have also become a dangerous tool for spreading misinformation and fake news causing severe damages to the life and property in the society. It has become imperative to find ways to distinguish authentic news from the misinformation on the internet to avoid social unrest and negative consequences. The current study makes the stance detection in news possible by extending the state-of-the-art COVAX reality dataset through the addition of two key features of a news body and a news summary. In addition, our study contributes a novel method comprising three different text summarization techniques (LEAD-3, sequence-to-sequence (Seq2Seq), and Bidirectional Encoder Representations from Transformers (BERT)) to authenticate the online news content. The BERT model consistently outperformed the other two models by achieving the best precision value (89.57%), Recall (91.84%), F-measure (90.64%) and BERT score (0.833). Our findings also prove the feasibility of using the transformer-based model, ELECTRA-Large, to combat the spread of fake news with 89.07% accuracy.

Keywords: BERT, Text Summarization, Fake News Detection, Stance Detection, NLP

Introduction

Fake news has spread widely in last few years and has posed a significant challenge to societies worldwide. The rapid spread of misinformation has undermined trust in traditional and non-traditional news sources. As a result, there is a need to combat the side effects of fake news propagation through the use of machine learning and NLP techniques (Chen, Lai, & Lian, 2023) (Ruffo et al., 2023) (Ahmed et al., 2023) (Devarajan et al., 2023).

The text summary is the task of condensing text content into a precise version while emphasizing the main points. Human-written summaries can produce pieces of text that are not in the source. Therefore, there is a need to create an AI model that can summarize text as close to a human-written summary as possible. In the context of fake news detection, abstractive text summarization techniques are employed to generate informative and coherent summaries that capture the essence of the original news articles (Onah, Pang, & El-Haj, 2022).

Our research focuses on the development and evaluation of an ML-based model for the fake news detection. The model leverages a diverse set of NLP features, such as linguistic, stylistic, contextual, and abstractive summaries. The objective is to construct a predictive model that can discern between authentic and fake news articles, thus providing a foundation for the development of more robust and efficient fake news detection systems (Elyassami et al., 2022).

This research paper focuses on a comprehensive approach to fake news detection by integrating multiple techniques, including BERT, stance detection, and sequence-to-sequence attention models. The primary objective is to construct a powerful model that can effectively identify fake news articles by extracting essential information, analyzing the generated summaries, and detecting the stance portrayed in both the news title and the AI-generated news summary.

The BERT model is employed to enhance the fake news detection process. This model leverages the extracted summaries and utilizes transformer-based mechanisms to capture the interdependencies between different parts of the text. By considering contextual relationships and dependencies, the model can discern rules that indicate fake news and improve the accuracy of the detection process.

Objectives

Our study aims to achieve the following goals:

- A new and all-encompassing approach for identification of fake news by enriched COVAX-Reality dataset.
- To create a model that uses text summarization with stance detection by utilizing BERT between the generated summary and the news headline.
- A valuable addition to current initiatives aimed at curbing the propagation of fake news by offering a detailed examination of both context of news stories and the content.
- An extensive evaluation of three techniques - Seq2Seq, LEAD-3 and BERT, using BERT Score and ROUGE metrics, along with an analysis of precision, F1-score, and recall.

Literature Review

We explore the text summarization techniques related work in this section. We focus on various deep learning techniques that are used to summarize the text. Furthermore, it also highlights the techniques used for stance detection and summary evaluation. (Saha Joy et al., 2022) focus on detecting COVID-19 related fake news.

The authors used (SVM) Support Vector Machine (Zhang, 2001) and achieved a high F1-score of 93.32% for classification of fake and real news. The authors also utilized ensemble methods and a heuristic algorithm to detect fake COVID-19 tweets, incorporating username and link domains. Multiple models, such as Bidirectional Encoder Representations from Transformers (BERT), Convolutional Neural Network (CNN) (Gu et al., 2018), and Long Short-Term Memory (LSTM) (Hochreiter & Schmidhuber, 1997) were employed in detecting fake news using the Covid 19 Fake News dataset. Language models like BERT (Devlin & Chang, 2018), ALBERT (Soricut & Lan, 2019), and XL- NET (Yang et al., 2019) were utilized in another re- search study to detect COVID-19 fake news.

The authors emphasized the importance of detecting misinformation related to global epidemics and proposed five different models, with RoBERT (Liu et al., 2019) performing the best; It achieved precision of 97%, recall of 99%, and F1 scores of 98%. The model was able to achieve 99% precision, 96% recall, and 98% F1 scores for the fake class. Additionally, a hybrid model combining BERT and ALBERT showed promising results (Hande et al., 2021)].

(Petrescu et al., 2023) introduced the EDSA-Ensemble model, a novel approach that uses event detection with sentiment analysis for sentiments extraction from social networks. The model integrates network analysis and machine learning, bridging the gap between two isolated communities – NLP and the network analysis. The evaluation of sentiment analysis algorithms indicates that employing an ensemble model with multiple models and text preprocessing techniques improves sentiment detection results. The proposed EDSA- Ensemble model outperforms individual models and accurately determines the overall sentiment of an event. Future work involves enhancing the model with additional Sentiment Analysis models based on CNN and Generative Adversarial Networks (GAN) (Creswell et al., 2018), and expanding the text preprocessing module with various word and transformer embedding methods, e.g., FastText (fastText, 2019), GloVe (Global Vectors for Word Representation) (Pennington, Socher, & Manning, 2014), to identify and mitigate hate speech and misinformation spread in communities.

(Suryavardan et al., 2023) focus on the publication of a substantial real-world dataset that includes both text and image inputs, aiming to advance machine learning approaches for multimodal fact verification. The authors conduct data analysis and release baseline models to highlight the challenges and potential for improvement in this field. The study suggests enriching the dataset with the reasoning behind the classification of fake news and exploring the use of synthetic data to add complexity to the refute category. The baseline results demonstrate the effectiveness of using a Vision Transformer (ViT) (Zlatkova, Nakov, & Koychev, 2019) for extracting visual features and a Sentence-BERT (SBERT) (Jindal et al., 2019) for textual features, achieving a F1 score of 0.6499. The comparison of different models highlights the significance of images in the task and the improvement obtained by using the ViT model over the ResNet (Singhal et al., 2019) model.

In research related to text summarization and its evaluation techniques, (Liu et al., 2020) discuss the differences between two types of summaries: human-written and computer-generated. It emphasizes the importance of creative ability in computer-generated summaries to make them more similar to human-written summaries. The effectiveness of deep learning models, particularly in text summarization tasks, is heavily influenced by the quality, structure, and representativeness of the datasets used for training and evaluation. The lack of abstractive datasets in Chinese long text summarization is highlighted, with the existing dataset being extractive in nature, leading to the generation of unoriginal summaries. The authors propose a new dataset called CLTS+ that addresses these limitations by paraphrasing reference summaries and correcting factual inconsistencies. They also present an inherent metric based on co-occurrence words to assess the dataset quality. The authors summarize their contributions, including the creation of the CLTS+ dataset and an analysis of its proper- ties, as well as the evaluation of baselines and out-of-domain data. The CLTS+ dataset shows promise in improving the creative ability of models and can serve as a benchmark for automatic text summarization in Chinese.

(Hsu et al., 2018) provide an overview of related research works in neural network-based extractive summarization and abstractive summarization. Extractive summarization methods utilize neural networks to select sentences based on sentence vectors, while abstractive summarization methods use attention-based encoders to generate summaries. The paper also discusses the challenges of extractive methods in terms of readability and highlights various techniques employed in abstractive methods, such as pointer networks and attention constraints. Additionally, the paper introduces the concept of hierarchical attention and proposes a unified model that combines the strengths of both abstractive and extractive approaches. The model incorporates sentence-level attention from extractive methods and word-level attention from abstractive methods. An inconsistency loss function is introduced to enhance the cooperation between these two levels of attention. Experimental results demonstrate that the proposed model achieves superior results in terms of ROUGE scores, informative-ness, and readability compared to other baselines on the Daily Mail/CNN dataset.

The authors in (Islam et al., 2021) address the critical challenge of fake news detection on social media by proposing a novel three-step approach that leverages natural language processing and machine learning techniques. Firstly, it involves stance detection to assess the alignment of the news article's title with its content. Then it checks the credibility of the author. Finally, numerous machine learning algorithms, including decision random forest, trees, support vector machine (SVM) and logistic regression are employed to classify news as real or fake. The SVM algorithm yielded the highest accuracy of 93.15%. It also serves as a groundwork for future research in the domain fake news detection, as it offers a unique three-step approach that has not been explored previously. Additionally, the research suggests the potential development of a commercial tool for tagging news as fake and providing credibility ratings.

Another similar approach (Qaiser et al., 2023) on fake news detection discusses the growing problem of misinformation and fake news in the digital age, emphasizing its impact on various aspects of society. The research aims to use three machine learning algorithms (Passive Aggressive Classifier (PAC) (Crammer et al., 2006), Naïve Bayes(NB) (Xu, 2018), SVM), three deep learning models (Recurrent Neural Net- work (RNN), CNN, Bidirectional LSTM (Liu, Sun, Lin, & Wang, 2016)) and transformer based models (ELECTRA-SMALL, BERT, and ELECTRA-LARGE (Clark, Luong, Le, & Manning, 2020))

are used to analyze the data set's performance. The proposed methodology involves three main phases: dataset extension, data pre-processing, and an inference engine. The dataset is enhanced with real-time news, combining fake and real news samples. Data pre-processing includes removing special characters, stop words, and lemmatization to improve data quality. The inference engine uses various machine learning, deep learning models, and transformer-based models to classify news articles as fake or real. Results indicate that the ELECTRA LARGE model out- performs other deep learning methods, achieving an impressive 85.11% accuracy.

Table 1: Structured Literature Review

Authors	Focus Area	Techniques/Models Used	Dataset/Source	Key Findings
Saha Joy et al., 2022	Fake news detection (COVID-19)	SVM, ensemble methods, heuristic algorithms, CNN, BERT, LSTM, RoBERTa, ALBERT, XLNet	COVID-19 Fake News Dataset	RoBERTa achieved 98% F1, hybrid BERT+ALBERT also promising
Petrescu et al., 2023	Sentiment analysis from social networks	EDSA-Ensemble model (event detection + sentiment), CNN, GAN, FastText, GloVe	Not specified	EDSA-Ensemble outperforms individual models; future work to improve hate speech detection
Suryavardan et al., 2023	Multimodal fact verification	Vision Transformer (ViT), SBERT, ResNet	New real-world dataset (text + image)	ViT + SBERT yielded better macro F1 (0.6499) than ResNet; visual input is crucial
Liu et al., 2020	Chinese text summarization	Dataset: CLTS+; co-occurrence-based metric for evaluation	CLTS+	Dataset improves creativity and factual correctness in long-text summarization
Hsu et al., 2018	Extractive & abstractive summarization	Unified model: sentence-level + word-level attention, pointer networks, hierarchical attention	CNN/Daily Mail	Proposed model outperforms baselines in ROUGE, readability, informativeness
Islam et al., 2021	Fake news detection (social media)	Three-step NLP + ML approach: stance detection, author credibility, SVM, decision trees, RF, LR	Not specified	SVM highest accuracy (93.15%); proposed commercial tool for fake news tagging
Qaiser et al., 2023	Fake news detection	ML: PAC, Naïve Bayes, SVM; DL: RNN, CNN, BiLSTM; Transformers: BERT, ELECTRA SMALL/LARGE	Real-time enhanced dataset	ELECTRA LARGE achieved 85.11% accuracy; best among all models
Wang, 2017	Text summarization techniques & evaluation	Graph-based (TextRank, LexRank), ML (SVM, clustering), DL (RNN, CNN, LSTM), Transformer (BART, PEGASUS, LongT5), GANs	Review paper	Comprehensive taxonomy and technical breakdown of ATS approaches; challenges in hallucination, evaluation metrics, and future directions

Table 1 summarizes the current related literature review for the research. In addition, this (Wang, 2017) comprehensive survey explores the evolution and methodologies of automatic text summarization (ATS), focusing on extractive, abstractive, and hybrid techniques. It categorizes summarization based on document type, summary style, domain, and language, and presents traditional and modern methods like graph-based models (TextRank, LexRank), ML algorithms (SVM, clustering), and deep learning models (LSTM, RNN, CNN).

Abstractive summarization approaches such as Seq2Seq, GANs, PEGASUS, LongT5, and BART are discussed in-depth, along with evaluation techniques including ROUGE, BERT Score, and cosine similarity. The paper highlights challenges like resource consumption, hallucinations in generative models, and the need for improved creativity and semantic coherence.

Materials and Methods

The suggested model described this module that was developed for training and tuning data using the extended version of the COVAX-reality dataset

In Figure 5, the primary elements of the architecture are shown, and the comprehensive description of the components are as follows:

1. COVAX-Reality dataset enrichment
2. Text summarization
3. Stance detection using text similarity
4. News Detection

DOI: 10.3844/jcssp.2025.2570.2580

	B	C	D	E	F	G	H	I	J
1	Statement	Link	Date	Label	Body	Summary			
2	Second shots of COVID-19 vaccine c	https://globalnews.ca/news/756441	08-Jan-20	TRUE	Some New Brunsw	Some New Brunswickers will receive the			
3	Which province is winning the COV	https://globalnews.ca/news/770684	20-Mar-20	TRUE	The COVID-19 vacc	The COVID-19 vaccine rollout kicked off i			
4	University of Guelph's COVID-19 va	https://globalnews.ca/news/697337	22-Mar-20	TRUE	The University of C	The University of Guelph says its researcl			
5	An ad by Merck pharma tells peopl	https://www.boomlive.in/world/fa	15-Jul-20	FALSE	Facebook posts sh	Facebook posts appear to show a billboai			
6	Don't wait to vaccinate kids during	https://globalnews.ca/news/68717S	27-Apr-20	TRUE	Canadian resident	The Canadian Paediatrics Society "strong			
7	The existence of a canine coronavi	https://www.politifact.com/factche	01-May-20	FALSE	There is currently i	The vaccine is for canine coronavirus dise			
8	mRNA vaccines are capable of alter	https://www.snopes.com/fact-chec	10-Dec-20	FALSE	In mid-November,	In mid-November, a lengthy bit of viral a			
9	Due to the large number of people	https://www.politifact.com/factche	20-May-20	FALSE	Unfounded fears a	A recent Facebook post alleges that Ame			
10	How 'vaccine hesitancy' became a t	https://globalnews.ca/news/51173S	02-Apr-20	TRUE	Sabrina Bacchus is	Two-thirds of parents believe vaccinatio			
11	White House blocks new coronavir	https://globalnews.ca/news/738041	05-Oct-20	TRUE	The White House h	The White House has blocked new FDA g			
12	Many people every year, sometime	https://www.politifact.com/factche	06-Oct-20	FALSE	President Donald T	In the U.S. so far this year, about 210,000			
13	US billionaire Bill Gates was caught	https://www.rappler.com/newsbre	15-Oct-20	FALSE	US billionaire Bill (	US billionaire Bill Gates was caught on ca			
14	Coronavirus vaccine: Pfizer may ask	https://globalnews.ca/news/740093	16-Oct-20	TRUE	With Americans, B	The ambitious initiative known as COVA)			
15	Image shows the total number of v	https://www.boomlive.in/world/90	09-Mar-20	FALSE	An image has beer	The image shows a child with syringes on			
16	Facebook user Lea Andoy Lugo post	https://www.rappler.com/newsbre	19-Mar-20	FALSE	There is no vaccine	There is no vaccine or cure for the diseas			
17	Margaret Keenan, the first recipien	https://www.snopes.com/fact-chec	10-Dec-20	FALSE	In December 2020,	On Dec. 8, Margaret Keenan became the			

Fig. 1. A Sample of the dataset showing the features of the dataset.

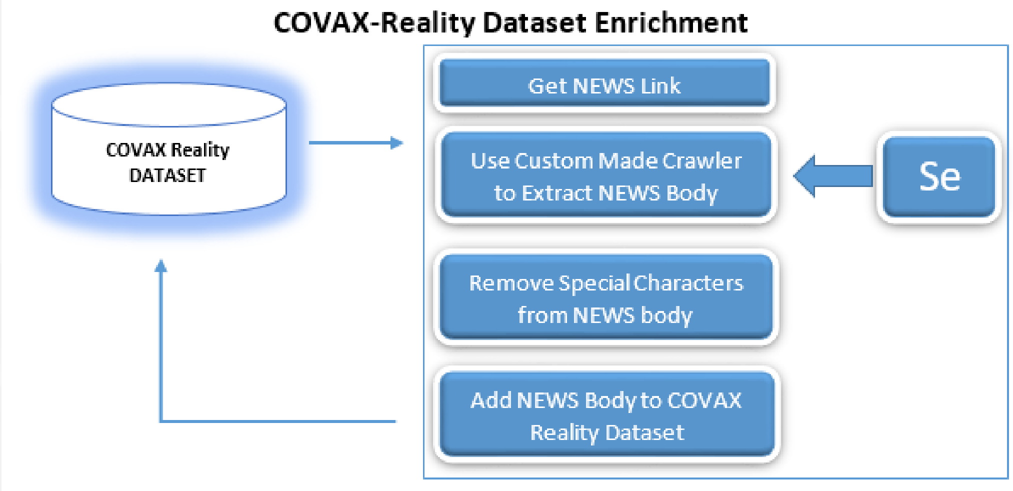

Fig. 2. Dataset Enhancement Process

COVAX-Reality Dataset Enrichment

Many datasets are available that contain data related to fake and real news (Wang, 2017; Shu, Mahudeswaran, Wang, Lee, & Liu, 2020). The dataset has proven effective in research applications, with many researchers achieving higher accuracy. The main problem of these datasets is that the majority of the dataset contains data about general news items. This research study mainly focuses on the COVID-19 vaccination news and very few datasets were available that deals with COVID-19 vaccination news. Due to this, the COVAX-Reality dataset is prepared, it specially provides the news about COVID-19 vaccination. A novel data set, COVAX-Reality (Qaiser et al., 2023) was chosen for this work. The dataset includes the following five features: "Id," "Title," "Link," "Author," and "Label." There are 2050 records in the dataset, 1210 of which are classified as legitimate news, while the rest of the records are false news. This dataset is dedicated to only COVID-19 vaccine-related news. All the news records were crawled from various trusted resources, such as PolitiFact, KESQ, Forbes, Snopes, Rappler, Boomlive, USA Today, the New York Times, CNN, etc. The news statements in the dataset were manually labeled as fake or real news. The fact-checking websites such as PolitiFact and Snopes was used to mainly collect fake news and the real news were mainly collected from the well-known news platforms like BBC, CNN, Rappler, and New York Times.

Table 2. Dataset Overview

Feature	Description
ID	A unique identifier allocated
Statement	News title
Link	URL of article
Date	Date of advertisement of news
Body	Description of the news
Summary	Summary of news article
Label	Tag of the news article

DOI: 10.3844/jcssp.2025.2570.2580

To enhance the quality and diversity of our original dataset, i.e., COVAX-Reality, it is enhanced by adding two additional features. One feature is the body of the news statement, and another feature is the summary of the news body. The description of the dataset is depicted in Table 2.

A subset of the few records in the dataset is showcased in Figure 1. The news crawler was created to extract the body of each headline from the news link. The extended version of COVAX- Reality enables us to perform stance detection, which is the first basic step in fake news detection.

In 2016, an online challenge named as Fake news challenge was initiated. The main motive of this challenge was to find new methods and approaches to detect the falsehood news. Stance detection was identified as the basic step to detect the news is fake or real. Stance detection is a technique which utilizes cosine similarity and finds whether news headline and body are both relevant. Figure 2 demonstrates the detailed process of dataset enhancement. Moreover, to transform the text of the news body into a summary, several text models for summarization such as (Seq2Seq with Attention, and BERT), (Baseline (Lead-3), were applied. The BERT model provided the incomparable results among all the summarization models. Properties of the fake news and real news are enlisted in the table 3.

Table 3. E-COVAX Reality Dataset Properties

News Property	Description
Total Count	2050
Real Count	1210
Fake Count	840
Real News Sources	BBC, CNN, Rappler, New York Times, etc.
Fake News Sources	PolitiFact, Snopes and etc.

Text Summarization

Language processing still has a lot to learn about summarizing lengthy texts. Automatic summary is critical since people have less patience and time to read lengthy texts. Large materials like stories, journal papers, news pieces, and even bigger texts like books may all be summarized automatically, which has important uses in this area. Extractive (Belwal, Rai, & Gupta, 2023) and Abstractive (Kouris, Alexandridis, & Stafylopatis, 2022) are two broad categories that might be used to group together existing summarization techniques (Allahyari et al., 2017).

To make text summarization on news articles, three different methods have been adopted, which are outlined:

1. Lead-3: Lead-3 (Moratanch & Chitrakala, 2017) uses extractive approach to make summary. The first three phrases in a document are chosen as the summary in the Lead-3 technique of summarizing. The Lead-3 baseline employs the first few phrases of a text as the summary since it thinks that they provide the most crucial information. This summarization model is easy to implement. Moreover, it does not require high computational resources.
2. Seq2Seq: Seq2Seq (Sutskever, Vinyals, & Le, 2014) follows an abstractive summarization model. In the sequence-to-sequence framework, encoder and decoder are the two components of recurrent neural network (RNN) (Sherstinsky, 2020). The RNN encoder which carries a single-layer bidirectional LSTM unit, reads the input sequence token by token and creates a series of hidden states that encoder which represents the input. The decoder's hidden states are created sequentially, one at a time by RNN decoder, resulting in the output sequence that serves as the summary.
3. BERT: BERT (Vaswani et al., 2017) can be used as an ex- tractive and abstractive summarization model. In this work, BERT is an abstractive model for summarization. Following are the steps BERT uses to summarize the news body, steps are depicted in Fig. 3.

a. Text preprocessing: By using Sentence Piece text preprocessing has been done. In this step, text is further divided into BERT-compatible tokens.

b. Input Preparation: This step often adds special tokens, such as the [CLS] (classification) token at the start of the input and the [SEP] (separator) tokens to indicate the division between text segments. This step converts the input data into the appropriate format required by the BERT model.

c. Input Encoding: Encoded representations are generated by passing the preprocessed input through the BERT model, which creates contextualized embedding for each token based on its surrounding context. For additional processing, you can acquire the [CLS] token's final hid- den states or those of other pertinent tokens.

d. Summarization: This step produces an abstract summary by inputting the encoded representations from the BERT encoder into a transformer-based decoder. Each summary token is auto regressively generated by the decoder based on the encoded representation.

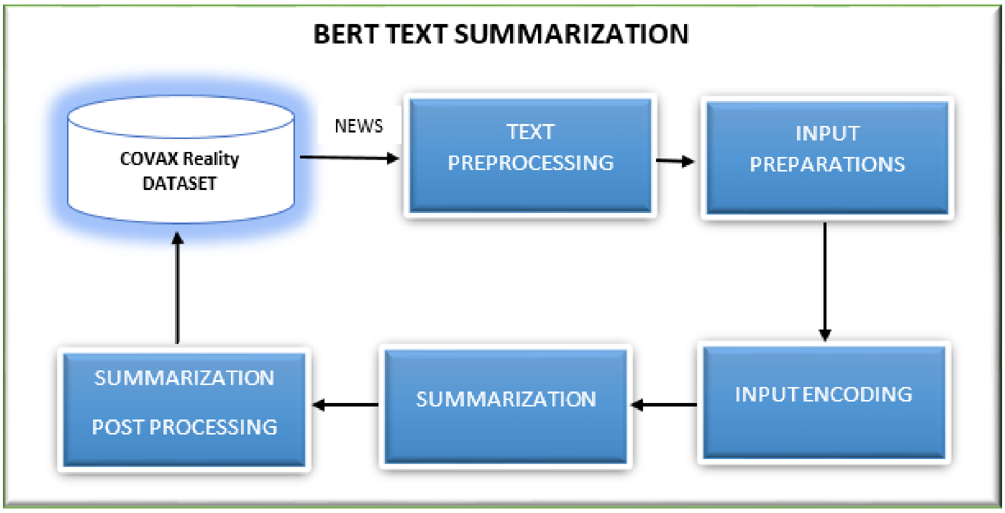

Fig. 3. Complete process of BERT text summarization

Table 4.

Models	Precision	Recall	F-Score
BERT	89.573%	91.844%	90.645%
Seq2Seq	82.544%	80.365%	81.444%
Lead-3	55.356%	51.914%	53.593%

e. Summarization Post-Process: This step provides the final version of the generated summary. If the summary contains any un- wanted special tokens, it eliminates all those tokens. Moreover, length constraints are imposed in this step.

Stance Detection via Text Similarity

In this module, it is determined whether the summary and the news headline are similar. To assess their relevance, the BERT model is utilized, as it is commonly employed to measure contextual similarity between two different texts.

In this module, it is determined whether the summary and the news headline are similar. To assess their relevance, the BERT model is utilized, as it is commonly employed to measure contextual similarity between two different texts.

If the news body summary and headline are not identical, one can move on to the next module, which is news detection; if not, the model identifies the news under scrutiny as fake news.

News Detection

To identify and detect the news in fake or real categories. A Transformer based ELECTRA-Large (Clark, Luong, Le, & Manning, 2020) model is used. As mentioned in the introduction, this research work is an extension of our previous work, in prior research, several different models were applied to the COVAX-Reality dataset. Among all the models we tested, the ELECTRA-Large transformer-based model provided the best accuracy and precision score. So, the obvious choice for the current study is ELECTRA-Large Model.

Evaluation Results

The evaluation of this approach is divided into two segments, news classification evaluation results and text summarization evaluation results.

Text Summarization Evaluation Results

In the initial phase, an evaluation was made between the news body text and the summary, whether the summarized text is accurate or not. For this, ROUGE metric (Lin, 2004) and Bert score (Zhang, Kishore, Wu, Weinberger, & Artzi, 2019) are used to measure the quality of the summary that the system has produced. To implement the Rouge metric, the following steps have been followed:

- Tokenization, removal of stop words, and punctuation cleaning were applied to both the news body and its corresponding summary text.
- The ROUGE score was calculated for each pair of news body and summary.
- This study specifically utilizes the ROUGE-1 metric for performance evaluation.
- Finally, the ROUGE score (average) across all made news bodies and summaries was computed to assess overall effectiveness.

The Rouge metric result is calculated and shown in Table 4 and BERT Score is also calculated for the evaluation of the text summary module. In Table 5, It is depicted that BERT gives 0.8335 BERT score which is the best result among all three models. After evaluating the text summary using two different matrices, it is concluded that out of the three techniques, BERT produces the most observable results.

Table 5. BERT Score

Models	Bert Score
BERT	0.8335
Seq2Seq	0.7215
Lead-3	0.5203

Result of News Classification

In the final phase, an evaluation is conducted to categorize the news as either fake or real. To assess the impact of enriched features (new body and summary) ablation study was conducted and the following observations have been made:

Impact of News Body: Addition of the news body has significantly made an impact on model performance.

Impact of News Summary: This feature has also made some improvement in the model performance but not as impactful as the news body.

Overall impact of news body and summary: The inclusion of these two features significantly improves the overall performance of the model, demonstrating that incorporating both the news body and summary provides a valuable contribution to this research.

To visually represent the model's performance, a confusion matrix for the ELECTRA-Large model was generated, with the corresponding confusion matrix is displayed in Figure 4. This matrix displays the true negatives, true positives, false negatives, and false positives, providing insight into the model's classification accuracy. Formulas used for calculations are:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$$

$$Precision = \frac{TP}{TP+FP}$$

$$Recall = \frac{TP}{TP+FN}$$

$$F1\ Score = \frac{2\ Precision \times Recall}{Precision+Recall}$$

The results indicate that ELECTRA-Large achieved a precision of 88.47%, an accuracy of 89.07%, a recall of 88.20%, and an F1-score of 87.45%. Table 6 shows the results, while the complete framework of the suggested model is presented in Figure 5.

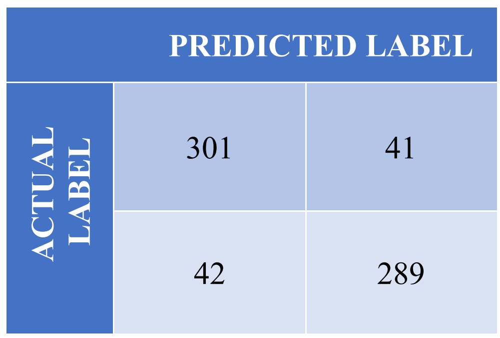

Fig. 4. Confusion Matrix

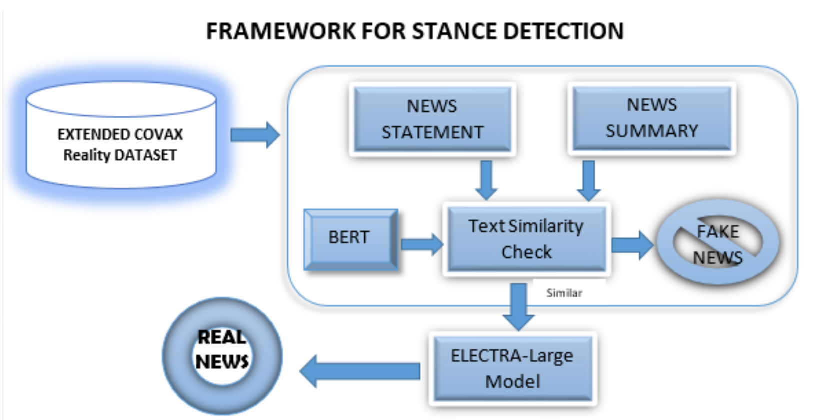

Fig. 5. Framework of Revealing Fake News

Table 6. The ablation study results (Enriched COVAX Reality dataset)

Feature used	Accuracy	Precision	Recall	F1 Score
Statement	85.110%	86.152%	83.585%	84.892%
Statement + Body	86.751%	87.253%	85.884%	86.463%
Statement + Summary	85.264%	86.644%	84.782%	85.374%
Statement + Body + Summary	89.075%	88.468%	88.192%	88.456%

Conclusion

This study extends the COVAX-Reality dataset through the integration of two novel features: the complete news body and its corresponding summary. These additions are intended to facilitate the task of stance detection, which constitutes a critical preliminary phase in the identification of misinformation. To this end, three distinct text summarization techniques were systematically evaluated to generate the summaries from the news content. Among all the summarization models we tested, the performance of BERT model was consistently better than others, producing the highest BERT score. Therefore, by using the BERT model, a news summary has been generated. Afterwards, comparisons have been made between the news summary and the news headline. If both are different, news will be declared fake news; otherwise, news will be classified as real and proceed towards the next step of the architecture, which is news detection.

The proposed approach did not address one of the important aspects which is the fake news explainability (De Magistris, Russo, Roma, Starczewski, & Napoli, 2022), that is, Why the news is classified as counterfeit remains a critical question. In the future, the FNEC model can be enhanced and made more interpretable by developing an explainable version, enabling a better understanding of the factors leading to the detection of fake news.

Acknowledgment

We extend our sincere gratitude to IQRA University, Karachi, Pakistan, for granting us access to the computer laboratories throughout the course of this study. This research was conducted in full compliance with the university's ethical guidelines for academic research. We also wish to express our appreciation to all the staff members whose support and cooperation were invaluable during this study.

Funding Information

The authors have not received any financial support or funding to report.

Data Availability Statement

The data can be made available on request.

Author's Contributions

Saad Ahmed: Provided essential guidance and oversight throughout the project. Designed the research plan, organized the study, coordinated data analysis, and contributed to the writing of the manuscript.

Asma Qaiser: Participated in all experiments, coordinated data analysis, implemented the research plan, and contributed to the writing of the manuscript.

Sidrah Abdullah: Designed the research plan, organized the study, participated in all experiments, coordinated data analysis, and contributed to the writing of the manuscript.

Muhammad Shoaib Siddiqui: Managed the administrative aspects of the project. Coordinated the data analysis and contributed to the writing of the manuscript.

Ethics

This manuscript represents original research conducted by the authors. The lead author affirms that all co-authors have thoroughly reviewed the content and have provided their full approval, with no ethical concerns raised.

References

Ahmed, S., Qaiser, A., Shahzad, J., Ali, M., & Khan, N. (2023). Detecting and analyzing deceptive information in news articles: A study using a dataset of misleading content. *International Journal of Emerging Engineering and Technology, 2*(1), 52–56.

Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). Text summarization techniques: A brief survey. *arXiv preprint arXiv:1707.02268*.

Belwal, R. C., Rai, S., & Gupta, A. (2023). Extractive text summarization using clustering-based topic modeling. *Soft Computing, 27*(7), 3965–3982. https://doi.org/10.1007/s00500-022-07409-5

Chen, M.-Y., Lai, Y.-W., & Lian, J.-W. (2023). Using deep learning models to detect fake news about COVID-19. *ACM Transactions on Internet Technology, 23*(2), 1–23. https://doi.org/10.1145/3579854

Clark, K., Luong, M.-T., Le, Q. V., & Manning, C. D. (2020). ELECTRA: Pre-training text encoders as discriminators rather than generators. *arXiv preprint arXiv:2003.10555*.

Crammer, K., Dekel, O., Keshet, J., Shalev-Shwartz, S., & Singer, Y. (2006). Online passive-aggressive algorithms. *Journal of Machine Learning Research, 7*, 551–585.

Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine, 35*(1), 53–65. https://doi.org/10.1109/MSP.2017.2765202

De Magistris, G., Russo, S., Roma, P., Starczewski, J. T., & Napoli, C. (2022). An explainable fake news detector based on named entity recognition and stance classification applied to COVID-19. *Information, 13*(3), 137. https://doi.org/10.3390/info13030137

Devarajan, G. G., Nagarajan, S. M., Amanullah, S. I., Mary, S. S. A., & Bashir, A. K. (2023). AI-assisted deep NLP-based approach for prediction of fake news from social media users. *IEEE Transactions on Computational Social Systems*.

Devlin, J., & Chang, M. (2018). Open sourcing BERT: State-of-the-art pre-training for natural language processing. *Google Research Blog*. https://ai.googleblog.com/2018/11/open-sourcing-bert-state-of-art-pre.html

Elyassami, S., Alseiari, S., ALZaabi, M., Hashem, A., & Aljahoori, N. (2022). Fake news detection using ensemble learning and machine learning algorithms. In *Combating fake news with computational intelligence techniques* (pp. 149–162). Springer. https://doi.org/10.1007/978-3-030-90055-7_8

fastText. (2019). *fastText blog*. https://fasttext.cc/blog/

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., & Cai, J. (2018). Recent advances in convolutional neural networks. *Pattern Recognition, 77*, 354–377. https://doi.org/10.1016/j.patcog.2017.10.013

Hande, A., Puranik, K., Priyadharshini, R., Thavareesan, S., & Chakravarthi, B. R. (2021). Evaluating pretrained transformer-based models for COVID-19 fake news detection. In *Proceedings of the 5th International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 766–772). IEEE. https://doi.org/10.1109/ICCMC51019.2021.9418394

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Hsu, W.-T., Lin, C.-K., Lee, M.-Y., Min, K., Tang, J., & Sun, M. (2018). A unified model for extractive and abstractive summarization using inconsistency loss. *arXiv preprint arXiv:1805.06266*.

Islam, N., Shaikh, A., Qaiser, A., Asiri, Y., Almakdi, S., Sulaiman, A., Moazzam, V., & Babar, S. A. (2021). Ternion: An autonomous model for fake news detection. *Applied Sciences, 11*(19), Article 19. https://doi.org/10.3390/app11198929

Jindal, S., Sood, R., Singh, R., Vatsa, M., Tanmoy, & Chakraborty, T. (2019). NewsBag: A multimodal benchmark dataset for fake news detection. *arXiv preprint arXiv:1905.10295*.

Kouris, P., Alexandridis, G., & Stafylopatis, A. (2022). Abstractive text summarization based on deep learning and semantic content generalization. *Information Processing & Management, 59*(1), 102790. https://doi.org/10.1016/j.ipm.2021.102790

Lin, C.-Y. (2004). ROUGE: A package for automatic evaluation of summaries. In *Proceedings of the Workshop on Text Summarization Branches Out* (pp. 74–81).

Liu, X., Zhang, C., Chen, X., Cao, Y., & Li, J. (2020). CLTS: A new Chinese long text summarization dataset. In *CCF International Conference on Natural Language Processing and Chinese Computing* (pp. 531–542). Springer. https://doi.org/10.1007/978-3-030-60450-9_44

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: An optimized method for pretraining self-supervised NLP systems. *arXiv preprint arXiv:1907.11692*.

Liu, Y., Sun, C., Lin, L., & Wang, X. (2016). Learning natural language inference using bidirectional LSTM model and inner-attention. *arXiv preprint arXiv:1605.09090*.

Moratanch, N., & Chitrakala, S. (2017). A survey on extractive text summarization. In *2017 International Conference on Computer, Communication and Signal Processing (ICCCSP)*. IEEE. https://doi.org/10.1109/ICCCSP.2017.7944061

Onah, D. F., Pang, E. L., & El-Haj, M. (2022). A data-driven latent semantic analysis for automatic text summarization using LDA topic modelling. In *Proceedings of the IEEE International Conference on Big Data* (pp. 2771–2780). IEEE.

Pennington, J., Socher, R., & Manning, C. D. (2014). GloVe: Global vectors for word representation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 1532–1543). https://doi.org/10.3115/v1/D14-1162

Petrescu, A., Truica, C.-O., Apostol, E.-S., & Paschke, A. (2023). EDSA-Ensemble: An event detection sentiment analysis ensemble architecture. *arXiv preprint arXiv:2301.12805*.

DOI: 10.3844/jcssp.2025.2570.2580

Qaiser, A., Hina, S., Kazi, A. K., Ahmed, S., & Asif, R. (2023). Fake News Encoder Classifier (FNEC) for online published news related to COVID-19 vaccines. *Intelligent Automation & Soft Computing, 37*(1).

Ruffo, G., Semeraro, A., Giachanou, A., & Rosso, P. (2023). Studying fake news spreading, polarisation dynamics, and manipulation by bots: A tale of networks and language. *Computer Science Review, 47*, 100531. https://doi.org/10.1016/j.cosrev.2022.100531

Saha Joy, S. K., Dofadar, D. F., Khan, R. H., Ahmed, M. S., & Rahman, R. (2022). A comparative study on COVID-19 fake news detection using different transformer-based models. *arXiv preprint arXiv:2208.XXXXX*.

Sherstinsky, A. (2020). Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena, 404*, 132306. https://doi.org/10.1016/j.physd.2019.132306

Shu, K., Mahudeswaran, D., Wang, S., Lee, D., & Liu, H. (2020). FakeNewsNet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media. *Big Data, 8*(3), 171–188. https://doi.org/10.1089/big.2020.0062

Singhal, S., Shah, R. R., Chakraborty, T., Kumaraguru, P., & Satoh, S. (2019). SpotFake: A multimodal framework for fake news detection. In *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)* (pp. 39–47). https://doi.org/10.1109/BigMM.2019.00044

Soricut, R., & Lan, Z. (2019). ALBERT: A lite BERT for self-supervised learning of language representations. *Google Research Blog*. https://ai.googleblog.com/2019/12/albert-lite-bert-for-self-supervised.html

Suryavardan, S., Mishra, S., Patwa, P., Chakraborty, M., Rani, A., Reganti, A., Chadha, A., Das, A., Sheth, A., & Chinnakotla, M. (2023). Factify 2: A multimodal fake news and satire news dataset. *arXiv preprint arXiv:2304.03897*.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems, 27*.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. In *Advances in Neural Information Processing Systems, 30*.

Wang, W. Y. (2017). “LIAR, LIAR pants on fire”: A new benchmark dataset for fake news detection. *arXiv preprint arXiv:1705.00648*.

Xu, S. (2018). Bayesian Naïve Bayes classifiers to text classification. *Journal of Information Science, 44*(1), 48–59. https://doi.org/10.1177/0165551516677946

Yang, Z., Dai, Z., Yang, Y., Carbonell, J., Salakhutdinov, R. R., & Le, Q. V. (2019). XLNet: Generalized autoregressive pretraining for language understanding. In *Advances in Neural Information Processing Systems, 32*.

Zhang, T. (2001). An introduction to support vector machines and other kernel-based learning methods. *AI Magazine, 22*(2), 103.

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., & Artzi, Y. (2019). BERTScore: Evaluating text generation with BERT. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3369–3383). https://doi.org/10.18653/v1/D19-1382

Zlatkova, D., Nakov, P., & Koychev, I. (2019). Fact-checking meets fauxtography: Verifying claims about images. *arXiv preprint arXiv:1908.11722*.