Fake News Encoder Classifier (FNEC) for Online Published News Related to COVID-19 Vaccines

Asma Qaiser[1], Saman Hina[1], Abdul Karim Kazi[1,*], Saad Ahmed[2] and Raheela Asif[3]

[1]Department of Computer Science and Information Technology, NED University of Engineering and Technology, Karachi, 75270, Pakistan
[2]Department of Computer Science, Iqra University, Karachi, 76400, Pakistan
[3]Department of Software Engineering, NED University of Engineering and Technology, Karachi, 75270, Pakistan
*Corresponding Author: Abdul Karim Kazi. Email: karimkazi@neduet.edu.pk
Received: 12 October 2022; Accepted: 13 December 2022

Abstract: In the past few years, social media and online news platforms have played an essential role in distributing news content rapidly. Consequently. verification of the authenticity of news has become a major challenge. During the COVID-19 outbreak, misinformation and fake news were major sources of confusion and insecurity among the general public. In the first quarter of the year 2020, around 800 people died due to fake news relevant to COVID-19. The major goal of this research was to discover the best learning model for achieving high accuracy and performance. A novel case study of the Fake News Classification using ELECTRA model, which achieved 85.11% accuracy score, is thus reported in this manuscript. In addition to that, a new novel dataset called COVAX-Reality containing COVID-19 vaccine-related news has been contributed. Using the COVAX-Reality dataset, the performance of FNEC is compared to several traditional learning models i.e., Support Vector Machine (SVM), Naive Bayes (NB), Passive Aggressive Classifier (PAC), Long Short-Term Memory (LSTM), Bi-directional LSTM (Bi-LSTM) and Bi-directional Encoder Representations from Transformers (BERT). For the evaluation of FNEC, standard metrics (Precision, Recall, Accuracy, and F1-Score) were utilized.

Keywords: Deep learning; fake news detection; machine learning; transformer model; classification

1 Introduction

In this era of social media, proliferation of information and spread of news globally is faster than ever before. This can be attributed primarily to an ever-increasing internet accessibility. According to data updated in January 2021, the number of internet users globally is about 4.66 billion and it is increasing at an incredible rate [1]. Nowadays, online social platforms serve as the primary source for disseminating information which includes news, user perspectives and expressions as well [2,3]. The main reason for this is that online news platforms are less costly and easily accessible.

Due to these merits, users usually prefer to consume news from online social platforms rather than conventional media like newspapers and television [4]. In the days of pandemics, uprisings, economic distress, and worry among the world's population, digital news platforms have been effective in terms of predicting and quickly disseminating useful information about issues of prime importance, for instance, pandemic outbreak trends (possibility of active cases), etc., [5]. On the other hand, it is also a fact that these platforms have some disadvantages as well such as the propagation of misinformation and purposely manufactured false news [6]. In recent times, the rapid spread of fake news linked to COVID-19 in particular has been witnessed all over the world using gamified approach [7]. Some examples of such fake news headers are "Cocaine kills coronavirus" and "Putin released lions on the streets of Russia to enforce coronavirus lockdown" [8]. Given that fake news are highly likely to leave a negative impact on society and individuals, fake news detection has become a critical task and. However, automatic fake news detection is still a challenging task due to the vigorous nature of online news content.

In the context of the aforementioned aspects of fake news, the contributions of this proposed framework are as follows.

- A novel dataset named "COVAX-Reality" containing COVID-19 vaccine-related news has been contributed to the research community.
- An in-depth analysis of various learning models on the novel fake news dataset COVAX-Reality and their accuracy comparison.
- A unique multi-staged framework for the COVID-19 vaccine news classification.
- A novel methodology for verifying the validity of the COVID-19 vaccination news (FNEC).

The above-mentioned contributions were devised after observing the need for automatic identification and classification of fake news to facilitate users in finding authentic information. In the presented research, an automated framework namely FNEC (Fake News Encoder Classifier) is proposed. This framework utilized the transformer-based model, i.e., ELECTRA (Efficiency Learning an Encoder that Classifies Token Replacements Accurately) that resulted in enhancing the accuracy. To evaluate the efficiency, the proposed method FNEC was compared against multiple state-of-the-art techniques (i.e., SVM, NV, PAC, LSTM, Bi-LSTM, and BERT) on the basis of standard performance metrics.

2 Related Work

Faster communication and easy access to information has led to widespread usage of social media all over the world [9,10]. However, there are crucial issues arising such as, spread of false information that has the potential to create chaos in the lives of individuals and as well as the society as a whole. If the news or information is fake, it could lead to bigger problems that can cause irreversible damage. For instance, a fake news titlted as "Alcohol has a possible cure for COVID-19" caused hospitalizations and deaths and deaths in Iran [11]. Thus, different means of authenticating information is becoming increasingly valuable [12,13]. Due to the vast amounts of information floating over the internet globally, an automatic means of authenticating information is bound to be incredibly useful.

Many technology enthusiasts, researchers, and Artificial Intelligent (AI) pursuers have been working on this issue of preventing social media users from the influence of fake news [14]. Various researchers have already made a significant contribution by developing different solutions/systems to detect false news on social media platforms. In this regard, a group of researchers proposed a way to utilize a combination of news contents and user comments to detect the authenticity of target news

using "Recurrent Neural Networks (RNN)" and "Gated Recurrent Unit (GRU)" [15]. They were not only able to detect the fakeness in news, but also achieved high accuracy as well. In another study, a model based on character n-grams and word n-grams features was formulated to detect fake news in online news [16]. Furthermore, by utilizing various lengths of n-grams, they predicted fake news by implementing two different machine-learning algorithms and achieved 96% accuracy. A group of researchers addressed the innovative problem of using social context for false news detection and a systematic technique to model the tri-relationship between news pieces, users, and publishers at the same time. This framework was based on learning news feature representations and predicted false news that utilizes both user-news interactions and publisher-news linkages [17]. Some researchers identified fake news by using a multi-stage approach [18].

Another research reported news encoding by using Bi-LSTM and then extracting the entities in the news headlines. Some prior studies also used two deep learning models using Bi-LSTM and GRU [19]. From the features of the LIAR dataset, the statement feature contributed the most in achieving an accuracy of 89.80%. A recent study proposed an architecture that used three-layer architecture. The first layer detects the stance, the second layer verifies the credibility of an author by using Twitter rule-based approach. In the last layer, three machine learning algorithms were applied to the fake news dataset with 93% accuracy achieved by the SVM algorithm [20]. Similarly, another study merged knowledge engineering with machine learning and framed a model to identify fake news that provided a decent accuracy [21]. Other than the application of various supervised machine learning algorithms [22,23], deep learning models [24–26], and transformer learning models [27–29], have also been used for the identification of fake news using different datasets. Moreover, another group of researchers used 19 different learning algorithms (8 machine learning algorithms, 6 deep learning algorithms, and 5 transformer algorithms) on 3 different datasets. 'Roberta', which is a transformer model performed better than the rest of the algorithms and provided 96% accuracy on combined corpus and fake or real datasets [30].

In contrast with the different approaches reported in the literature review, this study made use of datasets that were focused on news from any domain. Specifically, in our research, we focused on COVID-19 vaccination-related news. In contrast to the rule-based method or the conventional machine learning methods, FNEC uses the ELECTRA model. In addition, a comparison has been done with the other transformer models. To the best of our knowledge and understanding, this approach has not been reported in the available literature.

3 Comparative Study of COVID-19 Vaccine Datasets

Many datasets are available that contain data in the form of fake news and as well as authentic news [31–33]. Many researchers have utilized such datasets in their research and achieved high accuracy. The main dilemma of those datasets is that the majority data is about generic news. This study primarily focuses on the COVID-19 vaccination and other vaccine-related news, for which a very few datasets are available. Due to this, the 'COVAX-Reality' dataset is made because it specifically provides the most recent COVID-19 vaccination and vaccine-related news. This point is further explained by considering the example of a dataset namely 'ANTi-Vax'. This dataset consisted of 15,000 tweets and was manually annotated to be used in a research study [34]. The ANTi-Vax dataset focused on social media news, such as Twitter, whereas the COVAX-Reality dataset that is used in the present study focuses on traditional news. Furthermore, in comparison to the ANTi-Vax dataset, COVAX-Reality contains the most recent news.

Another example of the related dataset is a multimodal labeled dataset called MMCOVAR that is comprised of text, temporal information, and images of the news and tweets belonging to the COVID-19 vaccine [35]. For the classification of the news dataset, a combination of ratings from two distinct ranking sites was used that provides high-precision labeling. In this study, they not only provided a novel framework to test and validate the data of the MMCOVAR dataset but also provided a methodology of stance detection, that helped to annotate the tweets with an accuracy of 91.9%. The MMCOVAR dataset is a multimodal dataset, but COVAX-Reality exclusively deals with news statements. Additionally, the MMCOVAR dataset only has data up to May 8, 2021, whereas COVAX-Reality provides news data up to January 22, 2022.

4 Methodology

This section presents our proposed model that was trained and validated on the COVAX-reality dataset. The comparison of the proposed model is made with a variety of supervised machine learning algorithms, using word embedding with different deep learning models and transformer models. The proposed architecture of FNEC is comprised of the following components: COVAX-Reality Data Collection, COVAX-Reality Data Pre-processing, Word Embedding, and Inference Engine as shown in Fig. 1.
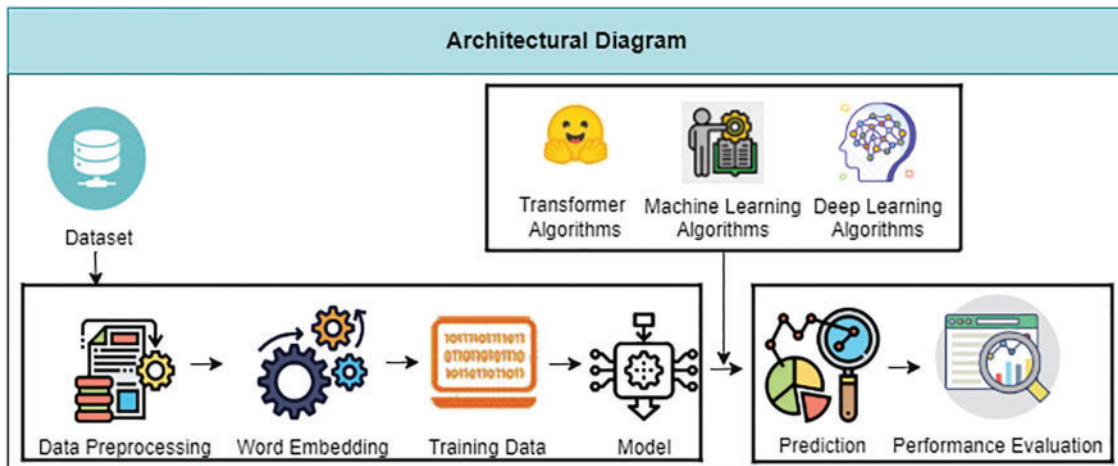

Figure 1: Architectural diagram of FNEC

4.1 COVAX-Reality Data Collection

As mentioned earlier, this study used a new dataset called the "COVAX-Reality" dataset containing a mixture of fake and authentic news in the English language. The novel dataset specifically focuses on the news related to the COVID-19 vaccination. To compile the dataset, numerous websites were crawled, including PolitiFact, Snopes, Boomlive, KESQ, Rappler, USA Today, Forbes, New York Times, CNN, etc. Fake news websites have been identified and crawled from the relevant websites [36]. A dataset containing real news was specifically crawled from trustworthy websites such as BBC, NewYork Times, CNN, and CNBC.

The COVAX-Reality dataset consists of five attributes named "ID", "Statement", "Link", "Date", and "Label". There were a total of 2050 news statements in the dataset that were crawled from 34 different publishers between 8th January 2020 and 25th January 2022. The "label" attribute is further

divided into two categories, one of which is labeled as "True" and the other as "False". 840 news statements belong to fake news and the rest belong to the authentic news.

The description of all features/attributes is tabulated in Table 1. Sample data from the COVAX-Reality dataset containing the fake and authentic news is presented in Table 2. This dataset news is further labeled into fake and authentic/real categories.

Table 1: COVAX-reality dataset description

Feature/Attribute	Description
ID	A distinct identifier for each news item
Statement	News headline
Link	The hyperlink to the news
Date	Published date of the news
Label	The label of the news

Table 2: COVAX-reality fake and real news examples

Statement	Label
One of the first nurses to receive the vaccine in AL is now dead.	False
COVID-19 patients died in Guinea after taking the trial vaccine.	False
Moderna, Aldevron Expand Partnership for mRNA COVID-19 Vaccine.	True
China coronavirus vaccine candidate appears safe for older people: preliminary results.	True

To analyze the importance of terms present in the dataset, 'Word clouds [37]' for both categories were generated to highlight a frequent group of words that are rendered with various sizes and lengths (as shown in Figs. 2, 3) 500 words were considered while creating a word cloud for the real and fake news. The bigger the word the more often the word appears in the news text statement. It can be observed that words like COVID, vaccine, people, vaccinated, etc., are more prominent in Fig. 2, indicating that these words are significantly contributing to the fake news statements, whereas Fig. 3 shows the important words in real news statements that contain different vaccine-related terms (AstraZeneca, Pfizer, dose, booster, health). This observation also shows that real news contains more relevant terms while fake news contains repetition of terms (Vaccine, Vaccinated, COVID, unvaccinated).

To obtain data records for true news, preference is given to trusted and well-known sources like BBC, KESQ, Reuters, USA Today, Forbes, New York Times, CNN, etc. The above-mentioned websites are the sites that always provide valid news so we can rely on the content of the website. PolitiFact is the only website that was automatically crawled by our custom-made web crawler written in python and the rest of the websites were manually crawled. An example of fact-checking website is shown in Fig. 4.

Fig. 5 depicts the news distribution concerning publishing dates starting from 8th January 2020 to 25th January 2022. It is observed that the number of news publications about COVID-19 vaccination has increased dramatically since December 2020. This timing corresponds to the availability of the COVID-19 vaccine in December, when the United Kingdom approved the Pfizer COVID-19 vaccine.
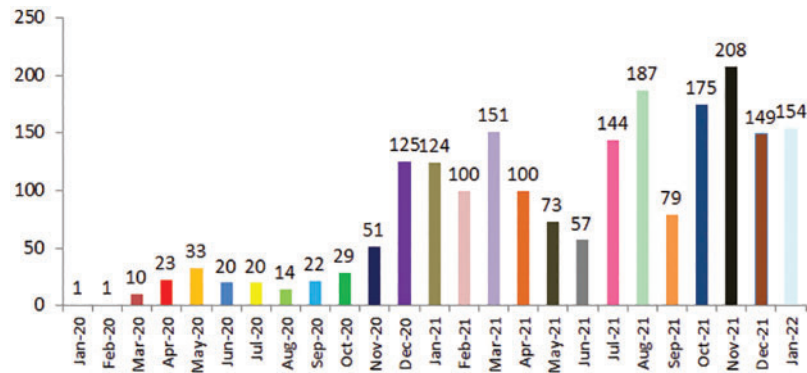
Figure 2: Word cloud representation of fake news

Figure 3: Word cloud representation of real news

Figure 4: An example of the fact-checking website (snopes.com)

Another rise in fake news was observed in November 2021, when in most countries the COVID-19 vaccination was available for teenage citizens.

Figure 5: A distribution of publishing date

Fig. 6a depicts the number of articles published in COVAX-Reality by various news outlets. The distribution of trustworthy publishers is shown with "Globalnews" and "PolitiFact" dominating, whereas, Fig. 6b depicts the distribution of articles by untrustworthy publishers, with "PolitiFact" dominating. PolitiFact website is dominating in both fake and real news because it is a fact-checking website and provides the label (fake or real) for each piece of news by cross-validating the factual information claimed in the news.
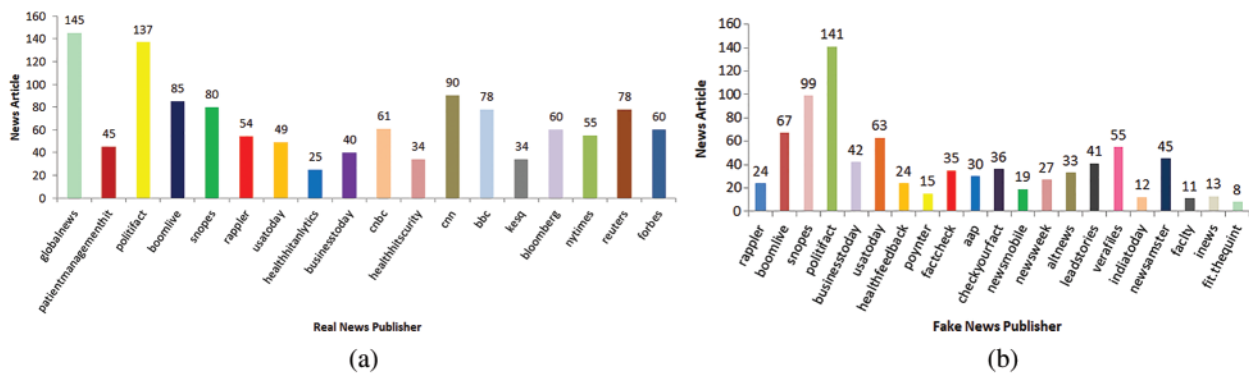
Figure 6: Distribution of (a) Real news and (b) Fake news publisher

4.2 COVAX-Reality Data Pre-Processing

To enhance efficiency, data pre-processing is an essential step that involves cleaning/normalization of the data before the implementation of methods. For the normalization of the dataset used in this research, certain steps were followed that eliminated the noise of the dataset, such as the removal of stop words, special characters, and punctuation from the text. Moreover, the text in different cases was converted into lowercase. Attributes of ID, link, and Date were also removed and only the statement attributes were used. After that, basic pre-processing modules (such as tokenization, sentence splitting, part-of-speech (POS) tagging, stop words elimination, and lemmatization) were applied, and named entity recognition (NER) module was also used to identify named entities within the dataset. The architectural representation of the steps followed in data pre-processing is shown in Fig. 7 and the steps are explained in detail.
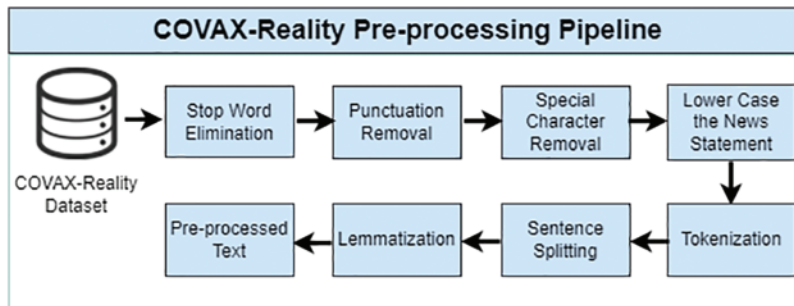
Figure 7: COVAX-Reality pre-processing pipeline

Step # 1: Stop word elimination: Words that add little to the meaning of a sentence are called stop words. They can be easily ignored because such words may carry little information about the text and may affect the efficacy of the model. For instance, words like are few examples of stop words may take significant processing time [38]. Therefore, after the removal of stop words, cleaned sample data is shown in Fig. 8.
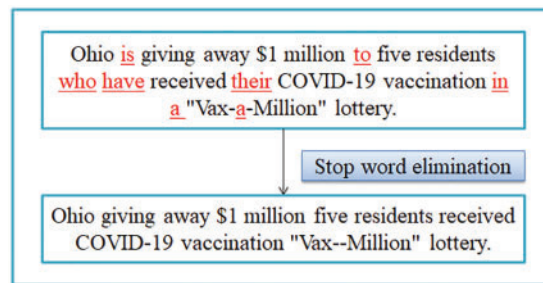

Figure 8: Sample output after stop word elimination

Step # 2: Punctuation marks removal: After the completion of step # 1, punctuation marks were also removed to enhance the performance. This was done because punctuation marks may or may not add meaning to the text. Sample output after the removal of punctuation marks is depicted in Fig. 9.
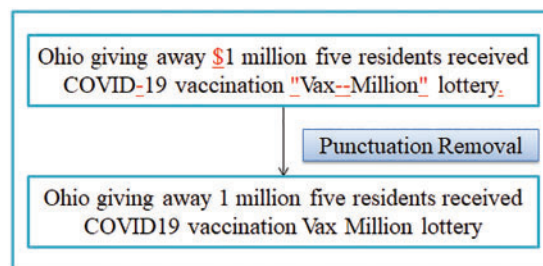

Figure 9: Sample output after punctuation removal

Step # 3: Removal of special characters: To collect the data for the COVAX-Reality dataset, a custom-made scrapper is made, which helps to scrape the news from the various sites. The code of the scraper is written in Python. When the data is scrapped using the custom-made scrapper, the data

contain some special characters (â€, â€™, â€œ, â€¦, etc.), which affect the readability of our COVAX-Reality news. To maintain the readability of the COVAX-Reality dataset, these special characters were removed from the data. These special characters do not contribute to the identification of fake news in the dataset. Therefore, cleaning the special characters would also improve the efficacy of the proposed model.

Step # 4: Conversion of text to lowercase: Lowercasing is the most typical pre-processing procedure. In this step text is converted to the lowercase. Sample output after the conversion of the text to lower-case is shown in Fig. 10.
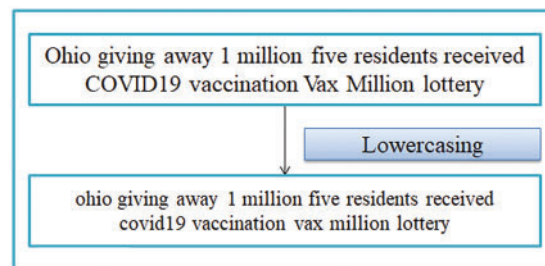

Figure 10: Sample output after conversion to lower-case

Step # 5: Language pre-processing: Before feeding the cleaned dataset to the proposed model, essential language pre-processing steps were applied. First, the dataset was tokenized using tokenizer and sentences were marked using a sentence splitter. Then, to match tokens with all the variants of targeted words, lemmatization was performed to get the root words from the dataset.

4.3 Word Embeddings

After the pre-processing of the text, the next phase is word embedding [39] to convert the words into vectors. Word embedding helps to learn models perform better on NLP applications [40]. In this research, some selected supervised machine learning algorithms and deep learning algorithms have utilized the word embedding technique, whereas transformer-based models used their embedding techniques.

4.4 Inference Engine

This phase is comprised of three categories (supervised learning, deep learning, and transformer models) of methods that have been used for the identification and classification of fake news, discussed Sections 4.4.1–4.4.3.

4.4.1 Supervised Machine Learning Models

As a baseline approach, traditional machine learning algorithms such as Naive Bayes (NB) [41], Passive Aggressive Classifier (PAC) [42], and Support Vector Machines (SVM) [43] have been applied for the classification. The best classification results were achieved by SVM (F1-Score: 79.21%), which are further discussed in Section 5.

4.4.2 Deep Learning Dense Models

In this approach, the following deep learning models were applied to evaluate if they could outperform classical supervised machine learning techniques. First model was Long Short-Term

Memory (LSTM), that [44] utilized two pre-trained word embeddings i.e., FastText and GLoVe [45]. LSTM network is a particular form of RNN [46], used to solve and learn long-term dependency problems. By using pre-trained word embeddings, the input data points are converted into word vectors. These word vectors become an input for the LSTM layer. With a 0.30 of dropout, we layered two LSTM layers one after the other. The LSTM has a dimension of 128, and the last time step output is used to represent input data points. The result of the final time step is fed into a dense layer for fake news detection. Second adopted model was Bidirectional Long Short-Term Memory (BiLSTM) [47], that was used with an embedding layer of vocabulary size of 5000, an input length of 20, and an embedding space of 40 dimensions. As an output layer activation function, 'Sigmoid' was used [48]. The optimization algorithm was an 'Adam' optimizer [49], and the model loss function was "binary cross-entropy". The model accuracy was evaluated by using the "accuracy" metric. '25 epochs' were used along with the batch size of '64' and the best model was selected.

4.4.3 Transformer Models

We have implemented the following transformer-based models; BERTBASE, ELECTRA SMALL, and ELECTRA LARGE. BERT was used as the baseline model (named as BERTBASE) in the selected transformer models. Among these three stated models; ELECTRA LARGE outperformed the machine learning and deep learning models and stands out with higher accuracy. We implemented these models by using the following hyperparameters, which are described in Table 3.

Table 3: BERT and ELECTRA model details

Model	Layers	Hidden size	Parameters
$BERT_{BASE}$	12	768	110 M
$ELECTRA_{SMALL}$	12	256	14 M
$ELECTRA_{LARGE}$	24	1024	335 M

As discussed earlier, the performance of ELECTRA LARGE is the best against the other learning models. Therefore, the ELECTRA LARGE is elaborated further in terms of parameters that outperformed the other learning models. ELECTRA LARGE introduces a new training objective to train large transformer models on raw data. Pre-training objective of ELECTRA LARGE is more efficient and leads to better performance than the masked language model (MLM) introduced in BERT. Moreover, ELECTRA LARGE replaced token detection tasks as shown in Fig. 11. It pre-trained the transformer model as a large discriminator predicting the tokens that have been replaced with the lower capacity of replaced tokens, which was more efficient than mass language modeling. This is because, it defined the loss in the entire input sequence and removed the masked token from the training objective. ELECTRA LARGE consists of two trained neural networks (generator G and discriminator D). The primary component of the generator is an encoder, which turns an input tokens sequence $a = [a_1..., a_n]$ into a sequence of contextualized vector representations $v(x) = [v_1,...,v_n]$. The generator generated a certain token using a SoftMax layer at position [50].

With a sigmoid output layer, the discriminator predicts whether the token comes from the data rather than the generating distribution for a particular location. To perform masked language modeling (MLM), the generator is trained. For instance, $a = [a_1, a_2, ..., a_n]$, first MLM masked $m = [m_1, m_k]$ by making a selection from a random set of positions. Typically, the position is chosen from 1 to n. A [MASK] token replaces the selected tokens, denoted as follows; $a^{masked} = REPLACE(a, m, [MASK])$

from the masked tokens. The training is provided to the discriminator to tell the difference between tokens in the data and tokens that have been substituted by a generator train the discriminator to anticipate which tokens in 'a^{corrupt}' match the original input 'a' by replacing the masked-out tokens with generator samples and creating a corrupted example. Model inputs are built in a formal way according to the following functions; $m_i \sim unif\{1,n\} for\, i = 1\, to\, k$; $\{\hat{a}\}_i \sim p_G\left(a_i | a^{masked}\right) for\, i \in m$; $a^{corrupt} = REPLACE(a, m, \hat{a})$ and the combined loss function for the generator and discriminator is presented in Eq. (1).

$$min_{\theta_G, \theta_D} \sum_{a \in X} L_{MLM}(a, \theta_G) + \lambda L_{Disc}(a, \theta_D) \tag{1}$$
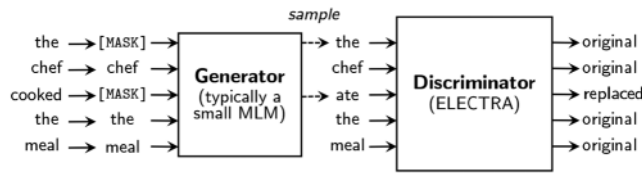
where 'X' is the raw text of a large corpus.

Figure 11: Replaced token detection overview [50]

5 Experimental Results

For evaluation purposes, the proposed architecture was implemented by using Python [51]. To get started with the experiment, the COVAX-Reality dataset was pre-processed using the Spacy [52]. Then, all the techniques of pre-processing were applied to the COVAX-Reality dataset (discussed in Section 4.2). After this, the text was transformed into vector form using a word embedding technique, and then three distinct machine learning algorithms were used as baseline approach: Passive Aggressive Classifier (PAC), Naive Bayes (NB), and Support Vector Machine (SVM). To summarize the classification performance of each supervised learning model, a confusion matrix was used in which True negative, True positive, False positive, and False negative values were analyzed [53]. The following are the definitions for each of these terms: True Positive (α): Legitimate news is predicted as legitimate; True Negative (β): illegitimate news is predicted as illegitimate news; False Positive (τ): illegitimate news is predicted as legitimate; False Negative (δ): legitimate news is predicted as illegitimate.

As mentioned earlier, the SVM classifier outperformed the other machine learning classification algorithms. The confusion matrix of SVM is shown in Fig. 12. As can be seen that '233' True Positives and '317' True Negatives were found and 81.84% accuracy was achieved. In the various scenarios, accuracy is not the best measure for assessing classification models. As a result, alternative metrics such as Recall (γ), Precision (ρ), and F1-score (η) should be used as standard performance metrics [54].

F1-Score of SVM is mathematically represented in Eq. (2).

$$F1 - Score\,(\eta) = \frac{2\rho\gamma}{\rho + \gamma} = \left(2 * \frac{(0.7766)\,(0.8090)}{0.7766 + 0.8090}\right) * 100 = 79.21\% \tag{2}$$

Fig. 13 shows the confusion matrix of Naive Bayes (NB). The NB classifier acquired an accuracy of 79.01%, a precision of 79.22%, a recall of 75.50%, and an F1-Score of 77.31%. Similarly, Fig. 14 shows the confusion matrix of the Passive Aggressive Classifier (PAC) which acquired an accuracy of 75.50%, a precision value of 77.66%, a recall value of 78.18%, and an F1-score value of 77.91%.
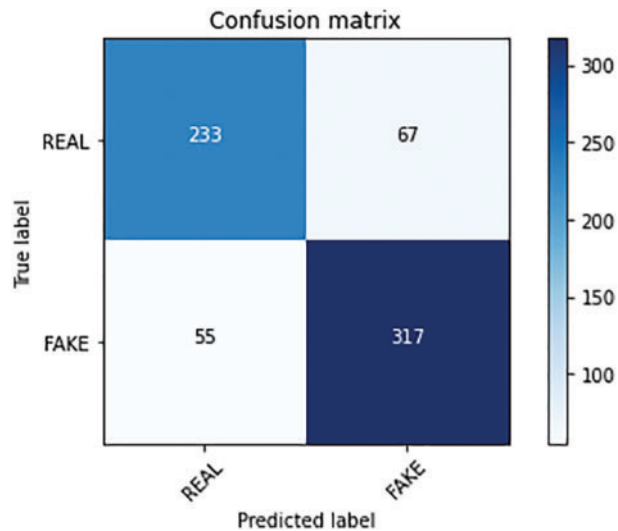
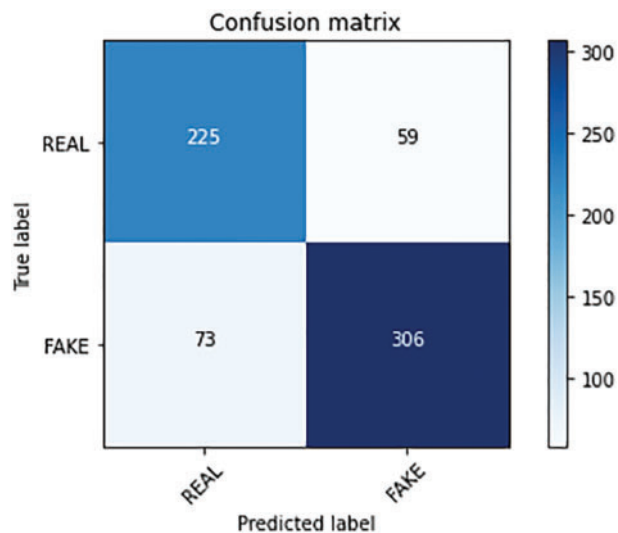**Figure 12:** SVM confusion matrix

Figure 13: NB confusion matrix

As discussed earlier, three deep learning models (LSTM + fastText, LSTM + GLoVe, and Bi-LSTM) were implemented and evaluated. BiLSTM performed well and provided 82.14% accurate results. After the implementation of transformer models, the results were compared after implementing all the classifiers. It was discovered that the ELECTRALARGE give the best accuracy and F1-score for the proposed FNEC against the mentioned classifiers. As tabulated in Table 4, the ELECTRALARGE acquired 85.11% of accuracy, 86.15% of precision, 83.58% of recall, and 84.85% of F1-score. This also shows the comparison of all the other applied methods, which are graphically represented in Figs. 15–18. When comparing the ELECTRALARGE to the second-best classifier, the ELECTRASMALL, the ELECTRALARGE clearly outperformed the ELECTRASMALL in terms of F1- score and accuracy. Accuracy and F1-score have been improved by 1.78% and 2.57%, respectively.
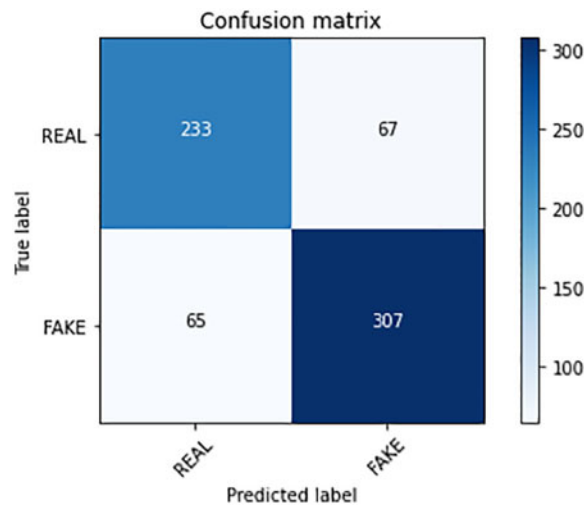
Figure 14: PAC confusion matrix

Table 4: Comparative analysis of performance for each learning model

| Model type	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Machine learning models	PAC	75.50	77.66	78.18	77.91
	NB	79.01	79.22	75.50	77.31
	SVM	81.84	77.66	80.90	79.21
Deep learning models	LSTM + fastText	79.55	79.65	74.27	76.86
	LSTM + GLoVe	80.05	78.57	74.82	76.64
	BiLSTM	82.14	81.23	78.55	79.87
Transformer models	$BERT_{BASE}$	80.95	76.21	78.93	77.54
	$ELECTRA_{SMALL}$	83.33	80.75	83.87	82.28
	$ELECTRA_{LARGE}$	85.11	86.15	83.58	84.85
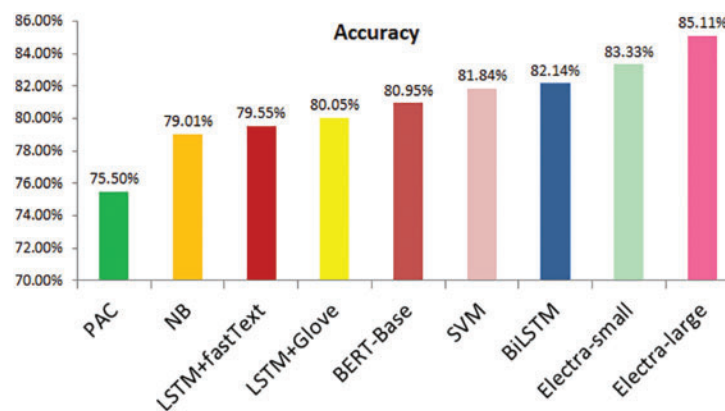
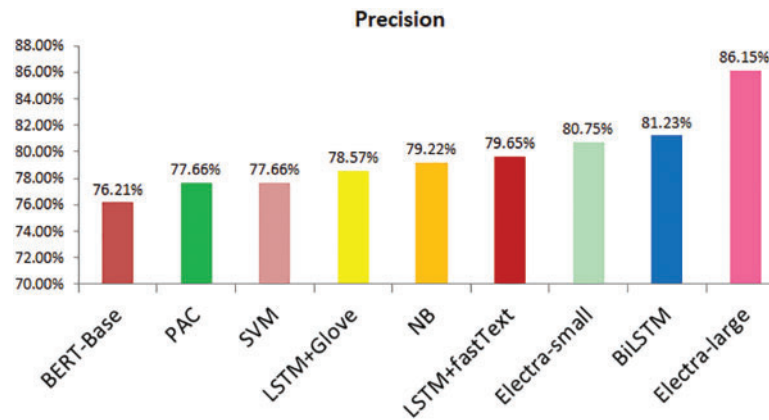**Figure 15:** Accuracy comparison attained by different evaluation models

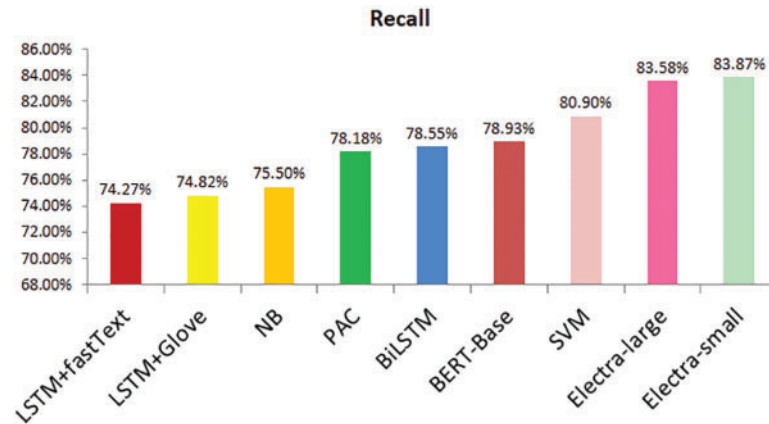**Figure 16:** Precision comparison attained by different evaluation models

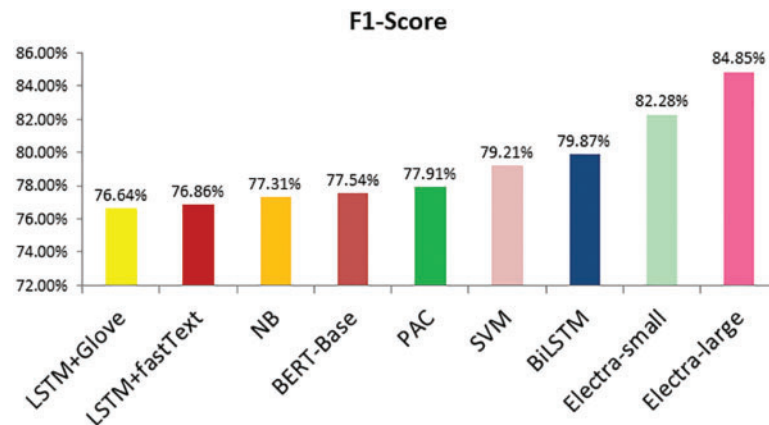**Figure 17:** Recall comparison attained by different evaluation models

Figure 18: F1-score comparison attained by different evaluation models

6 Conclusions and Future Work

Social media and online news platforms have been playing decisive roles in the distribution of news content at a lightning pace for the past few years. Among the consequences of the rapid and unruly spread of news is the propagation of unverified news. This research contributes a new model FNEC: Fake News Encoding Classifier to combat the ongoing COVID-19 crisis. The approach utilized numerous algorithms to model the novel dataset COVAX-reality, which is a new dataset consisting of Vaccination news of COVID-19. The ELECTRALARGE, which is a transformer-based model performed well and gave an accuracy of 85.11% and an F1-score of 84.85%.

The proposed approach can be further improved in the future. As an extension of the presented work, fake news explanation may aim at a detailed analysis of the news shared by news publishers as well as news shared on social media platforms. This may reveal answers to some critical questions [55], such as, "why is the news detected as counterfeit?". The FNEC model can be expanded and made more understandable in the future to detect fake news by making an explainable FNEC model. The suggested architecture can also be extended by adding one more feature in the COVAX-Reality dataset, that is, 'body of the news'. From a broader perspective, there is a need to deploy such a framework over the internet [56] to capture the authenticity of global news.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] S. Ghavifekr and S. Y. Wong, "Role of big data in education: Challenges and opportunities for the digital revolution in Malaysia," in *Handbook of Research on Big Data, Green Growth, and Technology Disruption in Asian Companies and Societies*. IGI Global: Pennsylvania, USA, pp. 22–37, 2022.

[2] S. Ahmed, S. Hina and R. Asif, "Detection of sentiment polarity of unstructured multi-language text from social media," *International Journal of Advanced Computer Science and Applications*, vol. 9, no. 7, pp. 199–203, 2018.

[3] S. Ahmed, S. Hina, E. Atwell and F. Ahmed, "Aspect based sentiment analysis framework using data from social media network," *International Journal of Computer Science and Network Security*, vol. 17, no. 7, pp. 100–105, 2017.

[4] O. B. Okunoye and A. E. Ibor, "Hybrid fake news detection technique with genetic search and deep learning," *Computers and Electrical Engineering*, vol. 103, no. 3, pp. 108344, 2022.

[5] S. Mohan, A. K. Solanki, H. K. Taluja, Anuradha, A. Singh *et al.,* "Predicting the impact of the third wave of COVID-19 in India using hybrid statistical machine learning models: A time series forecasting and sentiment analysis approach," *Computers in Biology and Medicine*, vol. 144, no. 10, pp. 105354, 2022.

[6] M. A. Strasser, P. J. Sumner and D. Meyer, "COVID-19 news consumption and distress in young people: A systematic review," *Journal of Affective Disorders*, vol. 300, pp. 481–491, 2022.

[7] L. Clever, D. Assenmacher, K. Müller, M. V. Seiler, D. M. Riehle *et al.,* "Fakeyou!-a gamified approach for building and evaluating resilience against fake news," in *Multidisciplinary Int. Symp. on Disinformation in Open Online Media*, Leiden, The Netherlands, Springer, pp. 218–232, 2020.

[8] A. Bovet and H. A. Makse, "Influence of fake news in Twitter during the 2016 US presidential election," *Nature communications*, vol. 10, no. 1, pp. 1–14, 2019.

[9] P. Boczkowski, E. Mitchelstein and M. Matassi, "Incidental news: How young people consume news on social media," in *Proc. of the 50th Hawaii int. conf. on system sciences (HICSS-50)*, Hilton Waikoloa Village, Hawaii, pp. 1785–1792, 2017.

[10] Y. K. Kim and K. L. Fingerman, "Daily social media use, social ties and emotional well-being in later life," *Journal of Social and Personal Relationships*, vol. 39, no. 6, pp. 1794–1813, 2022.

[11] P. Patwa, S. Sharma, S. Pykl, V. Guptha, G. Kumari *et al.,* "Fighting an infodemic: Covid-19 fake news dataset," in *Int. Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situation*, Springer, pp. 21–29, 2021. https://aaai.org/Conferences/AAAI-21/

[12] D. Orso, N. Federici, R. Copetti, L. Vetrugno and T. Bove, "Infodemic and the spread of fake news in the COVID-19-era," *European Journal of Emergency Medicine*, vol. 27, no. 5, pp. 327–328, 2020.

[13] S. Boberg, T. Quandt, T. Schatto-Eckrodt and L. Frischlich, "Pandemic populism: Facebook pages of alternative news media and the corona crisis-a computational content analysis," arXiv preprint arXiv:2004.02566, 2020.

[14] F. Olan, U. Jayawickrama, E. O. Arakpogun, J. Suklan and S. Liu, "Fake news on social media: The impact on society," *Information Systems Frontiers*, vol. 23, pp. 1387–3326, 9th January, 2022 (Online), 2022.

[15] K. Shu, L. Cui, S. Wang, D. Lee and H. Liu, "Defend: Explainable fake news detection," in *Proc. of the 25th ACM SIGKDD Int. Conf. on Knowledge Discovery & Data Mining*, Anchorage AK, USA, pp. 395–405, 2019.

[16] H. E. Wynne and Z. Z. Wint, "Content based fake news detection using n-gram models," in *Proc. of the 21st Int. Conf. on Information Integration and Web-based Applications & Services*, Munich, Germany, pp. 669–673, 2019.

[17] I. Ahmad, M. Yousaf, S. Yousaf and M. O. Ahmad, "Fake news detection using machine learning ensemble methods," *Complexity*, vol. 2020, no. 5, pp. 1–11, Article Id. 8885861, 2020.

[18] M. Mayank, S. Sharma and R. Sharma, "DEAP-FAKED: Knowledge graph based approach for fake news detection," arXiv preprint arXiv: 2107.10648, 2021.

[19] N. Aslam, I. Ullah Khan, F. S. Alotaibi, L. A. Aldaej and A. K. Aldubaikil, "Fake detect: A deep learning ensemble model for fake news detection," *Complexity*, vol. 2022, no. 4, pp. 1–8, Article Id. 5557784, 2021.

[20] N. Islam, A. Shaikh, A. Qaiser, Y. Asiri, S. Almakdi *et al.,* "Ternion: An autonomous model for fake news detection," *Applied Sciences*, vol. 11, no. 19, pp. 9292, 2021.

[21] V. Mehta and R. K. Mishra, "Machine learning based fake news detection on covid-19 tweets data," in *Proc. of Int. Conf. on Computational Intelligence and Data Engineering*, India, Springer, pp. 89–96, 2022.

[22] S. Elyassami, S. Alseiari, M. ALZaabi, A. Hashem and N. Aljahoori, "Fake news detection using ensemble learning and machine learning algorithms," in *Combating Fake News with Computational Intelligence Techniques*, Switzerland: Springer, pp. 149–162, 2022.

[23] S. Ahmed, K. Hinkelmann and F. Corradini, "Combining machine learning with knowledge engineering to detect fake news in social networks-a survey," arXiv preprint arXiv: 2201.08032, 2022.

[24] A. Govindaraju and J. Griffith, "Classifying fake and real neurally generated news," in *2021 Swedish Workshop on Data Science (SweDS)*, Växjö, Sweden, IEEE, pp. 1–7, 2021.

[25] H. Saleh, A. Alharbi and S. H. Alsamhi, "OPCNN-FAKE: Optimized convolutional neural network for fake news detection," *IEEE Access*, vol. 9, pp. 129471–129489, 2021.

[26] T. Chauhan and H. Palivela, "Optimization and improvement of fake news detection using deep learning approaches for societal benefit," *International Journal of Information Management Data Insights*, vol. 1, no. 2, pp. 100051, 2021.

[27] Y. Wang, Y. Zhang, X. Li and X. Yu, "COVID-19 Fake news detection using bidirectional encoder representations from transformers based models," arXiv preprint arXiv: 2109.14816, 2021.

[28] R. K. Kaliyar, A. Goswami and P. Narang, "FakeBERT: Fake news detection in social media with a BERT-based deep learning approach," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 11765–11788, 2021.

[29] L. J. Y. Flores and Y. Hao, "An adversarial benchmark for fake news detection models," arXiv preprint arXiv: 2201.00912, 2022.

[30] J. Y. Khan, M. T. I. Khondaker, S. Afroz, G. Uddin and A. Iqbal, "A benchmark study of machine learning models for online fake news detection," *Machine Learning with Applications*, vol. 4, no. 2, pp. 100032, 2021.

[31] W. Y. Wang, ""liar, liar pants on fire": A new benchmark dataset for fake news detection," arXiv preprint arXiv:1705.00648, 2017.

[32] K. Shu, D. Mahudeswaran, S. Wang, D. Lee and H. Liu, "Fakenewsnet: A data repository with news content, social context and spatiotemporal information for studying fake news on social media," *Big Data*, vol. 8, no. 3, pp. 171–188, 2020.

[33] J. Thorne, A. Vlachos, C. Christodoulopoulos and A. Mittal, "Fever: A large-scale dataset for fact extraction and verification," arXiv preprint arXiv: 1803.05355, 2018.

[34] K. Hayawi, S. Shahriar, M. A. Serhani, I. Taleb and S. S. Mathew, "ANTi-Vax: A novel twitter dataset for COVID-19 vaccine misinformation detection," *Public Health*, vol. 203, no. 5, pp. 23–30, 2022.

[35] M. Chen, X. Chu and K. Subbalakshmi, "MMCoVaR: multimodal COVID-19 vaccine focused data repository for fake news detection and a baseline architecture for classification," in *Proc. of the 2021 IEEE/ACM Int. Conf. on Advances in Social Networks Analysis and Mining*, Netherlands, pp. 31–38, 2021.

[36] Y. Tsfati, H. G. Boomgaarden, J. Strömbäck, R. Vliegenthart, A. Damstra *et al.,* "Causes and consequences of mainstream media dissemination of fake news: Literature review and synthesis," *Annals of the International Communication Association*, vol. 44, no. 2, pp. 157–173, 2020.

[37] L. Oesper, D. Merico, R. Isserlin and G. D. Bader, "WordCloud: A cytoscape plugin to create a visual semantic summary of networks," *Source Code for Biology and Medicine*, vol. 6, no. 1, pp. 1–4, 2011.

[38] Y. Nezu and T. Miura, "Extracting stopwords on social network service," *Information Modelling and Knowledge Bases XXXI*, vol. 321, pp. 59, 2020.

[39] S. Ghannay, B. Favre, Y. Esteve and N. Camelin, "Word embedding evaluation and combination," in *Proc. of the Tenth Int. Conf. on Language Resources and Evaluation (LREC'16)*, Portorož, Slovenia, pp. 300–305, 2016.

[40] K. Chowdhary, "Natural language processing," *Fundamentals of Artificial Intelligence*, Chapter 19, ISBN 978-81-322-3972-7 (ebook), New Delhi, India, pp. 603–649, 2020.

[41] S. Xu, "Bayesian naïve bayes classifiers to text classification," *Journal of Information Science*, vol. 44, no. 1, pp. 48–59, 2018.

[42] K. Crammer, O. Dekel, J. Keshet, S. Shalev-Shwartz and Y. Singer, "Online passive aggressive algorithms," *Journal of Machine Learning Research*, vol. 7, pp. 551–585, 2006.

[43] T. Zhang, "An introduction to support vector machines and other kernel-based learning methods," *AI Magazine*, vol. 22, no. 2, pp. 103, 2001.

[44] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[45] J. Pennington, R. Socher and C. D. Manning, "Glove: Global vectors for word representation," in *Proc. of the 2014 Conf. on Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, pp. 1532–1543, 2014.

[46] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Physica D (Nonlinear Phenomena)*, vol. 404, no. 8, pp. 132306, 2020.

[47] Y. Liu, C. Sun, L. Lin and X. Wang, "Learning natural language inference using bidirectional LSTM model and inner-attention," arXiv preprint arXiv: 1605.09090, 2016.

[48] A. Wanto, A. P. Windarto, D. Hartama and I. Parlina, "Use of binary sigmoid function and linear identity in artificial neural networks for forecasting population density," *IJISTECH (International Journal of Information System & Technology)*, vol. 1, no. 1, pp. 43–54, 2017.

[49] Z. Zhang, "Improved adam optimizer for deep neural networks," in *2018 IEEE/ACM 26th Int. Symp. on Quality of Service (IWQoS)*, Banff, Alberta, Canada, IEEE pp. 1–2, 2018.

[50] K. Clark, M.-T. Luong, Q. V. Le and C. D. Manning, "Electra: Pre-training text encoders as discriminators rather than generators," arXiv preprint arXiv:2003.10555, 2003.

[51] G. vanRossum, "Python reference manual," *Department of Computer Science [CS]*, no. R 9525, CWI, pp. 1–59, 1995. https://ir.cwi.nl/pub/5008

[52] M. Honnibal and I. Montani, "spaCy 2: Natural language understanding with bloom embeddings, convolutional neural networks and incremental parsing," *To Appear*, vol. 7, no. 1, pp. 411–420, 2017.

[53] C. Sammut and G. I. Webb, *Encyclopedia of machine learning*.NewYork, USA: Springer Science & Business Media, 2011.

[54] C. Goutte and E. Gaussier, "A probabilistic interpretation of precision, recall and F-score, with implication for evaluation," in *European conf. on Information Retrieval*, Santiago de Compostela, Spain, Springer, pp. 345–359, 2005.

[55] G. De Magistris, S. Russo, P. Roma, J. T. Starczewski and C. Napoli, "An explainable fake news detector based on named entity recognition and stance classification applied to COVID-19," *Information-an International Interdisciplinary Journal*, vol. 13, no. 3, pp. 137, 2022.

[56] A. K. Kazi, S. M. Khan and N. G. Haider, "Reliable group of vehicles (RGoV) in VANET," *IEEE Access*, vol. 9, pp. 111407–111416, 2021.