2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

# PLNCC: Leveraging New Data Features for Enhanced Accuracy of Fake News Detection

Keshopan Arunthavachelvan
*Toronto Metropolitan University*
Toronto, Canada
karunthavachelvan@torontomu.ca

Shaina Raza
*Vector Institute of Artificial Intelligence*
Toronto, Canada
shaina.raza@vectorinstitute.ai

Chen Ding
*Toronto Metropolitan University*
Toronto, Canada
cding@torontomu.ca

***Abstract*—The prominence of social media poses a significant threat to information integrity as the spread of fake news increases. It becomes imperative for online media outlets to develop effective strategies to mitigate the spread of fake news. In this research, the PLNCC dataset is introduced as an expansion of two state-of-the-art fake news datasets. The objective is to improve the classification of fake news by extracting additional linguistic and psychological features, as well as user comment data. In this work, a quantitative analysis of the linguistic and psychological features of fake news articles and related user comments is performed. The efficacy of the PLNCC dataset is demonstrated through rigorous evaluation, showcasing its performance against state-of-the-art benchmark datasets. The classification models running on this dataset achieved a significant performance improvement, up to 10%, when compared to the original two datasets.**

***Index Terms*—fake news detection, linguistic features, psychological features, user engagement, user comments, fake news dataset**

## I. Introduction and Background

In an era where social media has popularized the digital dissemination of information, the spread of fake information presents a substantial threat to maintaining the trust and integrity of information in modern society. As fake news becomes increasingly prominent, it is crucial for social media platforms to develop preventative measures to effectively recognize and mitigate fake news on their platforms.

Earlier fake news detection [1] primarily relied on the news article itself for classification by analyzing the news's content, writing style, and source credibility. However, researchers have started including users' engagements and propagation data (user profiles for news distribution and tracking) as part of the classification task. Previous studies [2] show that user engagement and news propagation provide valuable information for the classification of fake news. However, other work [3] have identified through analysis that identifying the propagation graph or news cascade [3] can be time-consuming, and this type of information may not be available if we want to detect fake news in the early stages. In this study, we employed a simplified and efficient approach for the early detection of fake news by taking into account the characteristics of the news itself as well as the accompanying users' comments data.



*ASONAM '23*, November 6-9, 2023, Kusadasi, Turkey

ACM ISBN 979-8-4007-0409-3/23/11…$15.00
http://dx.doi.org/10.1145/3625007.3627308

Cognitive Natural Language Processing (NLP) [4] is an active research area that integrates cognitive science, psychology, linguistics, and artificial intelligence to better understand and interpret human language. Its ability to process the grammatical and semantic structure of language as well as aspects and underlying sentiment of discourse has profound implications in distinguishing between real and fake news [5].

Several studies [6] have uncovered the role of psychological characteristics in differentiating between real and fake news. Other related studies [7] have highlighted the potential for extracting various sentiment-based characteristics from news data, including the article's text and user comments. These studies consider factors such as emotions, swear words, and social behaviour for detecting the veracity of the news. An in-depth analysis of these feature groups indicates that social behavioural traits, such as politeness, interpersonal conflict, moralization, pro-social behaviour and communication have elevated effects in fake and real news. These studies have also identified that fake news often exhibits emotional bias [8] and profanity, while truthful articles are more neutral. Positive and negative emotions, as well as tones of anxiety, sadness, and anger, are common traits found in fake news and are beneficial when incorporated with fake news classification [9].

Traditionally, text-based classification models primarily focus on analyzing the writing styles of news articles; however, recent studies show that incorporating user engagement into the detection task leads to better classification performance. Unlike news articles, user comments offer valuable insights into the public opinion regarding the article's subject, allowing classification models to assess the viewpoints of different individuals, as opposed to relying on only the author's viewpoint.

### A. Objectives

The main research objective for our work is to develop a Psycho-Linguistic News Content and Comments (PLNCC) dataset which enhances the accuracy of existing fake news classification models by leveraging the linguistic and psychological features extracted from the textual components of news articles and associated user engagement through user comments. To achieve this, we expand upon two state-of-the-art text-based fake news datasets, namely NELA [10] and Fakeddit [11], by incorporating linguistic and psychological features, along with user comments. The NELA-GT-2019 dataset consists of news articles collected from multiple sources during 2019. This data includes the article's text, news headline, metadata related to the publisher and time of publication, as well as a classification label for each news article. Fakeddit is a compiled list of user comments collected from different subreddits on Reddit, which relate to world news events occurring during the collection period. To improve the classification task of existing fake news models, we attempt to combine both the news content from the NELA dataset with the related user comments from the Fakeddit dataset.

By extracting linguistic features from the article's text and user comments, we obtain detailed information about the writing style and formality of the news. Real news articles are typically written in a more formal manner, while fake news often exhibits informal language and numerous abbreviations [12]. We also extract attributes from the article's text and user comments including disinformation-related traits and clickbait terms from the headline to identify fake news better [13]. Additionally, incorporating psychological features provides insights into the emotional aspects and biases present in the writing, which are common traits in fake news.

Including these features in fake news detection enables models to be trained to identify the distinctive linguistic and psychological patterns between fake and real news, leading to more accurate classification.

In this work, we answer the following research questions:

**RQ1** How do linguistic and psychological features affect the performance of fake news classification?

**RQ2** Does the consolidated dataset improve the performance of existing fake news classification?

**RQ3** Which linguistic and psychological feature groups exhibit the best performance?

We conducted a series of experiments to analyze the effectiveness of linguistic and psychological features in both news content and user comments for fake news research. The results of our experiments demonstrate a ten percent improvement in classification accuracy compared to baseline methods.

## II. Data Description

In this work, we merge the text-based content from NELA with the user comments from Fakeddit to create a unified dataset to incorporate news articles with user comments. Furthermore, we generate a diverse set of linguistic and psychological features extracted from the unified dataset, and offer these features directly in the PLNCC dataset for better classification. The dataset is available on our GitHub page[1].

### A. Data Pre-Processing

The NELA dataset consists of a balanced dataset of fake and real news articles collected from different news sources throughout 2019, consisting of the news headline, the article's text, and metadata related to the time and source of publication. Fakeddit sources its dataset over a ten-year span of Reddit user comments from several selectively chosen subreddits. The user comments, publication timestamps and metadata related to each comment comprise the dataset. The original news article headlines from NELA and Reddit post headlines from Fakeddit are matched to concatenate the two datasets. After concatenation, we extract the article headline and text, user comments, date and time of publication, and the binary classification label from NELA as features to create the pre-processed dataset. After concatenation, the PLNCC dataset becomes unbalanced with 70% fake and 30% real news articles. Formatting discrepancies between the news headlines from NELA and Fakeddit results in this change in distribution.

### B. Labeling

We inherit two-way labels from NELA dataset to classify all news articles in the dataset. Although Fakeddit provides two-, three-, and six-way labelling using the subreddit's credibility score, these labels are much broader compared to the news source level NELA uses. Thus, we utilize the labelling system from NELA for PLNCC, which obtains labelling based on the credibility of the publishing news source. Our final dataset contains zero to many user comments for each news article. One binary label is provided for the evaluation of the classification model for each article. A recent study by Raza and Ding [8] utilizes similar concatenation processes using NELA and Fakeddit, indicating that this form of dataset concatenation and labelling is effective for the classification of fake news.

## III. Methodology

Our particular contribution to fake news research is on improving the effectiveness of textual features extracted from the NELA and Fakeddit datasets for fake news classification. To achieve this, we utilize natural language processing (NLP) techniques, including parts of speech (POS), context-free grammar (CFG), disinformation-related attributes (DIA), clickbait-related attributes (CBA) [14] and BERT embedding [17] and leverage the Linguistic Inquiry and Word Count (LIWC) dictionary [18] to extract linguistic and psychological features from the input data.

We consider the task of identifying fake news as a binary classification problem, where the resulting output is either a real (0) or fake (1) integer. To evaluate the effectiveness of our dataset, we rely on standardized metrics [16] such as accuracy and F1 Score as evaluation metrics.

*1) Linguistic Features:* While most classification models use a common set of linguistic features for their detection task, fake news datasets commonly provide only the article's text content, causing classifiers to require manual extraction of these features each time. Our dataset attempts to simplify this process by including a set of commonly used linguistic features extracted from the article's text and user comments to improve the simplicity of classification models.

[1] https://github.com/kchelvan/PL-NCC

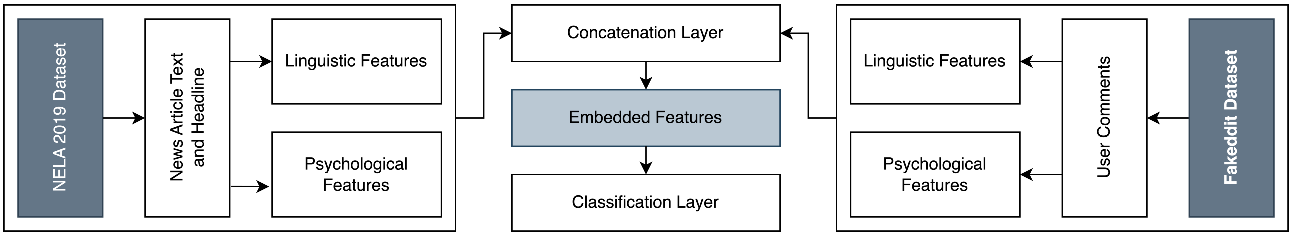


Fig. 1. Overview of PLNCC dataset

TABLE I
FAKE NEWS DETECTION ACCURACY ON OUR DATASET VS. ORIGINAL DATASET COMBINED OVER 5-FOLDS

| | **NELA Dataset** | | **Fakeddit Dataset** | | **PLNCC Dataset** Text-Only Input | | **PLNCC Dataset** *With Proposed Features* | | **Change in Accuracy** | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Model** | **Accuracy** | **F1 Score** | **Accuracy** | **F1 Score** | **Accuracy** | **F1 Score** | **Accuracy** | **F1 Score** | **Accuracy** | **F1 Score** |
| BERT | 0.920 | 0.942 | 0.789 | 0.847 | 0.868 | 0.908 | 0.972 | 0.980 | + 10.4% | + 7.2% |
| DistilBERT | 0.948 | 0.962 | 0.767 | 0.833 | 0.871 | 0.908 | 0.972 | 0.979 | + 10.1% | + 7.1% |
| MLP | 0.951 | 0.965 | 0.656 | 0.754 | 0.957 | 0.969 | 0.975 | 0.982 | + 1.8% | + 1.3% |
| XGBoost | 0.952 | 0.966 | 0.706 | 0.811 | 0.945 | 0.961 | 0.975 | 0.982 | + 3.0% | + 2.1% |
| CNN | 0.948 | 0.965 | 0.679 | 0.779 | 0.949 | 0.963 | 0.958 | 0.964 | + 0.9% | + 0.1% |
| DTC | 0.853 | 0.894 | 0.673 | 0.770 | 0.855 | 0.897 | 0.917 | 0.941 | + 6.2% | + 4.4% |

**Proposed Features** include User Comments, Linguistic and Psychological Features. BOW embedding of article headline and texts are the input to the models. The text-only input excludes user comments, linguistic, and psychological features.

Common features in most text-based classification models include POS and CFG; thus, these features are included in the dataset. Recent classifiers have extracted newer forms of embedding models such as BERT [17], [19] to aid in text-based classification. Thus, we extract BERT embeddings from the article's text, headline and user comments to include in the dataset. Additionally, the use of clickbait and disinformation-related attributes from the text has been explored [15] to aid in fake news classification. Disinformation traits are characteristics in the text's writing style which portray misinformation, such as discrepancy (terms which provide reasoning), causation (explanations), tentative (terms defining potential or conditions), insight (individual thoughts and knowledge), certitude (terms defining certainty), and differentiation (terms comparing variance). Clickbait attributes are common phrases used in news headlines, such as "this will blow your mind" and "can change your life", which are used to attract the attention of readers. These traits provide contrasting differences in fake and real news, which is beneficial in fake news classification. A higher score indicates that the feature is more likely to be real or fake news, respectively. We obtain a label for clickbait attributes by comparing the headline of the news article against a dictionary of forty-seven commonly used clickbait titles. If an article's headline contains any clickbait titles, it is assigned a score of one; otherwise, a score of zero is assigned.

*2) Psychological Features:* Recent studies [6], [7] in early fake news detection focus on different psychological features to improve the performance of the classification task. Existing classifiers using psychological features require manual extraction of these features from the article's text; however, our research indicates that all text-based classification models can benefit from the use of psychological features for their detection task. We classify psychological features into three feature groups: emotions, word toxicity, and social behavioural traits, which we include as part of our PLNCC dataset. Social behavioural traits such as morality, politeness, communication (addressing a topic or subject's stance), interpersonal conflict (conflict between two subjects), and pro-social behaviour (voluntary act to help others) show varying patterns in fake and real news. Additionally, we extract the negative and positive emotions within the article's text and user comments to assist with the classification task. With negative emotional features, tones such as anxiety, anger, and sadness can be extracted from the text and used for classifying fake news. Finally, word toxicity, such as the use of swear words is included as a feature in the PLNCC dataset.

Using the articles' text content and headlines, as well as user comment data from the combined NELA and Fakeddit datasets, we obtain the mentioned psychological features using the LIWC dictionary. LIWC utilizes TF-IDF weights to compare the frequency of words or phrases in the input text against the LIWC dictionary. LIWC then generates a numerical score out of one hundred for each psychological feature in the article's text and user comments. The resulting output indicates the percentage of terms in the input text included in the LIWC dictionary. Section IV illustrates how the patterns in feature scores differ between fake and real news for each individual feature. Additionally, all linguistic and psychological features in our dataset, excluding BERT embedding, are represented as a percentage score out of 100. We illustrate these patterns in Table II and perform an in-depth analysis in Section IV to study the effects of the proposed features and their scores in fake news classification.

TABLE II
FEATURE COMPARISON BETWEEN FAKE AND REAL NEWS (HIGHER AVERAGE SCORE MEANS MORE PROMINENT)

| | Article | | | | Comments | | | |
|---|---|---|---|---|---|---|---|---|
| | **Fake News** | | **Real News** | | **Fake News** | | **Real News** | |
| **Feature** | Avg | Max | Avg | Max | Avg | Max | Avg | Max |
| Clickbait | 0.00 | 1.00 | 0.02 | 1.00 | - | - | - | - |
| Insight | 1.74 | 7.71 | 2.16 | 18.2 | 1.58 | 33.3 | 1.23 | 50.0 |
| Causation | 1.39 | 5.98 | 1.19 | 11.1 | 1.09 | 20.0 | 0.80 | 25.0 |
| Discrep. | 0.99 | 9.52 | 1.23 | 10.0 | 1.26 | 40.0 | 0.85 | 25.0 |
| Tentat. | 1.35 | 10.3 | 1.68 | 9.09 | 1.89 | 33.3 | 1.11 | 33.3 |
| Certit. | 0.27 | 2.99 | 0.70 | 10.0 | 0.53 | 40.0 | 0.64 | 100 |
| Differ. | 2.47 | 7.94 | 2.01 | 20.0 | 2.38 | 22.2 | 1.56 | 25.0 |
| Toxicity | 0.01 | 2.04 | 0.27 | 10.5 | 0.58 | 100 | 0.40 | 25.0 |
| Pos Emo. | 0.29 | 7.14 | 0.50 | 13.3 | 0.52 | 33.3 | 0.42 | 20.0 |
| Neg Emo. | 0.43 | 8.89 | 0.67 | 8.33 | 0.80 | 50.0 | 0.51 | 25.0 |
| Pro Soc. | 0.61 | 6.06 | 0.48 | 7.14 | 0.35 | 20.0 | 0.23 | 50.0 |
| Polite. | 0.34 | 4.90 | 0.16 | 4.76 | 0.29 | 16.7 | 0.13 | 10.0 |
| Conflict | 0.65 | 9.09 | 0.61 | 11.8 | 0.27 | 25.0 | 0.18 | 9.09 |
| Moral. | 0.44 | 5.54 | 0.50 | 14.3 | 0.35 | 33.3 | 0.22 | 25.0 |
| Comm. | 2.70 | 14.3 | 1.56 | 11.1 | 1.22 | 25.0 | 0.74 | 50.0 |

**Avg.** represents the **Average Embedding Score** obtained for each feature. **Max.** represents the highest **Feature Embedding Score** obtained for each feature. **Clickbait attributes** relate to news headlines only, thus, do not have a comment score.

## IV. EXPERIMENTS AND RESULTS ON PLNCC DATASET

The PLNCC dataset is run against several state-of-the-art fake news classification models using NLTK, Keras, and Scikit-Learn. For the classification task, 70% of the dataset is used for training, while 30% is used to test. To minimize training loss, each model is trained for fifteen epochs and optimized using the Adam optimizer. The linguistic features extracted are embedded with a dimensionality size of 768, and the feature scores extracted using LIWC are used as is for the model's classification. Each model uses a ReLU activation layer, a softmax output layer, and utilizes the sparse categorical cross-entropy loss function. To create a holistic analysis of our dataset's effectiveness, we use baseline models, including multilayer perceptron models (MLP), XGBoost, convolutional neural networks (CNN) and decision tree classifiers (DTC). Recent studies [19], [20] have adapted BERT embedding to improve the performance of fake news classification; thus, we include BERT embedding models with the MLP classifier.

### A. Patterns of Linguistic and Psychological Features

Firstly, we analyze the effects of each linguistic and psychological feature in fake and real news and observe the following patterns. The results of our analysis are presented in Table II, where we illustrate the score distribution for each linguistic and psychological feature between real and fake news. Each feature in the table has a minimum value of zero.

In our analysis, we identify that fake news articles are more likely to have clickbait titles compared to real news articles, as there is a higher average of fake news articles with a higher feature score compared to real news. Through our analysis, user comments exhibit higher scores of disinformation-related attributes (DIA) in fake news than in real news, such as discrepancy; however, the article text has varying results for the same traits in both fake and real news. Differentiation in the article's text and headline is equally distributed between real and fake news, while certitude is much more frequent in real news than fake, as illustrated in Table II.

Recent work with cognitive NLP [6], [7] exploring emotions in fake news classification has identified that fake news is emotionally biased compared to real news. Analyzing negative emotional scores shows that emotions are more present in fake news, as indicated by the significantly higher scores in fake news compared to real news. Additionally, user comments in fake news show higher emotional scores, while real news articles show higher negative emotional scores. When analyzing real news articles which contain higher toxicity scores, we discover that specific news sources, such as the Onion, are recognized as reliable news sources; however, their writing style contains a substantial amount of offensive language in their headlines, such as "depressed monkey throwing sh*t at himself". Sources such as the Onion provide news articles based on current events; however, they are written in the form of satire. Although the content of the news source is truthful, these news articles are considered outliers in the data. Additionally, user comments with high levels of toxicity are often correlated with fake news. Analyzing the different patterns between the news article and user comments provides valuable insight into the importance of user comments in fake news classification, as different sets of information can be extracted from the two data types.

We compare different feature groups to examine the effects of linguistic and psychological features, as depicted in Table III. As we gradually include more of the proposed features in the classification task, the model's performance consistently improves, with an increase of up to twenty-five percent from the first experiment in the series. By incorporating user comments and linguistic and psychological features into our dataset, fake news classification models are better equipped to predict fake news compared to traditional text-based datasets.

### B. Effectiveness of User Comments

While articles represent the viewpoint of a single writer, user comments provide varying perspectives from the public. Negative emotions in comments indicate conflicting views on the article's topic, while positive emotions can indicate support. By examining user comments and word toxicity, we observe a correlation between an increase in toxic language in user comments and news articles that receive higher negative emotional scores from users. As fake news tends to drive controversial topics [1] to generate user engagement and promote news propagation, this can result in users expressing heightened emotional views about the news topic, resulting in increased toxic language in their comments. Our research also reveals that real news has a higher average score in negative emotions within the article content, as indicated in Table II. Since news sources cover topics such as natural disasters, violence, and politics, negative bias may be present in the writing, signifying a higher negative emotional score.

We examine the effects of various social behavioural traits in fake and real news, and present our results in Table II.

TABLE III
ABLATION STUDY USING ARTICLE TEXT AND USER COMMENTS

| Model | Accuracy | F1 Score |
|---|---|---|
| BOW Only | 0.723 | 0.778 |
| BERT Only | 0.868 | 0.908 |
| BERT + POS + CFG | 0.956 | 0.965 |
| BERT + DIA + CBA | 0.944 | 0.957 |
| Linguistic + Emotion | 0.963 | 0.960 |
| Linguistic + Social Behaviour | 0.965 | 0.952 |
| Linguistic + Swear Words | 0.968 | 0.964 |
| Linguistic + Psychological | 0.972 | 0.979 |

Traits such as morality and interpersonal conflict are more prevalent in the article content in fake news. Scores closer to 0.5 for morality indicate the news article is morally just, while higher scores for conflict indicate the news contains more conflict-related information. Additionally, pro-social behaviour, politeness and communication have elevated scores in user comments and the article's text in real news. A lower score for these features indicates that these features are more prominent in the news. Using these linguistic and psychological features and user comments provides valuable information in the classification task and improves the performance of existing detection models. Leveraging these patterns allows classifiers to better differentiate between fake and real news.

### C. Effectiveness of PLNCC Dataset Against Baseline Models

We compare the performance of various state-of-the-art baseline classifiers against the NELA, Fakeddit, and PLNCC datasets. We evaluate these models using only the article headline and text, and then include the proposed user comments, linguistic and psychological features, as shown in Table I. Results show the PLNCC dataset improves the performance of existing models by up to ten percent when the proposed features are incorporated during the model's classification. When using the same feature sets, we observe an improvement in performance by utilizing both article text and user comments from the PLNCC dataset compared to solely relying on NELA's article content or Fakeddit comment data.

## V. CONCLUSION

Existing fake news datasets provide only the text content of news articles, requiring models to extract various sets of linguistic and psychological features for fake news classification. However, this step increases the complexity of classifiers. We contribute to the field of fake news research by: 1) performing an in-depth analysis of the effects of different linguistic and psychological features in real and fake news; 2) creating a compiled dataset to incorporate user propagation data with text-based classification by including the articles' text content with related user comments; and 3) developing an improved fake news dataset which includes pre-embedded values of effective linguistic and psychological features to improve the efficacy of existing fake news classification models. We execute a series of experiments to demonstrate the effectiveness of the PLNCC dataset against state-of-the-art classifiers and illustrate the benefits of user comments, linguistic and psychological features in fake news research. The results of our experiments indicate the PLNCC dataset provides improvements to the performance of fake news classification models.

## ACKNOWLEDGMENT

This work is partially sponsored by Natural Science and Engineering Research Council of Canada (grant 2020-04760).

## REFERENCES

[1] B. D. Horne and S. Adali, 'This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news', in Eleventh international AAAI conference on web and social media, 2017.
[2] X. Zhou and R. Zafarani, 'Network-based fake news detection: A pattern-driven approach', ACM SIGKDD explorations newsletter, vol. 21, no. 2, pp. 48–60, 2019.
[3] F. Qian, C. Gong, K. Sharma, and Y. Liu, 'Neural user response generator: Fake news detection with collective user intelligence', in IJCAI International Joint Conference on Artificial Intelligence, 2018, vol. 2018-July, pp. 3834–3840.
[4] N. R. de Oliveira, P. S. Pisa, M. A. Lopez, D. S. V. de Medeiros, and D. M. Mattos, Identifying fake news on social networks based on natural language processing: trends and challenges, Information, 12(1) 38, 2021.
[5] A. Acerbi, "Cognitive attraction and online misinformation," Palgrave Communications, vol. 5, no. 1, 2019.
[6] X. Zhang, J. Cao, X. Li, Q. Sheng, L. Zhong, and K. Shu, 'Mining dual emotion for fake news detection', in Proceedings of the Web Conference 2021, 2021, pp. 3465–3476.
[7] C. Guo, J. Cao, X. Zhang, K. Shu, and M. Yu, Exploiting emotions for fake news detection on social media, arXiv preprint:1903. 01728, 2019.
[8] S. Raza and C. Ding, 'Fake news detection based on news content and social contexts: a transformer-based approach', International Journal of Data Science and Analytics, vol. 13, no. 4, pp. 335–362, 2022.
[9] S. Gaillard, Z. A. Oláh, S. Venmans, and M. Burke, "Countering the cognitive, linguistic, and psychological underpinnings behind susceptibility to fake news: A review of current literature with special focus on the role of age and digital literacy," Frontiers in Communication, vol. 6, p. 661801, 2021.
[10] M. Gruppi, B. D. Horne, and S. Adalı, 'NELA-GT-2019: A Large Multi-Labelled News Dataset for The Study of Misinformation in News Articles', arXiv [cs.CY]. 2020.
[11] K. Nakamura, S. Levy, and W. Y. Wang, 'r/fakeddit: A new multimodal benchmark dataset for fine-grained fake news detection', arXiv preprint arXiv:1911. 03854, 2019.
[12] G. Pennycook and D. G. Rand, "The psychology of fake news," Trends in cognitive sciences, vol. 25, no. 5, pp. 388–402, 2021.
[13] E. C. Tandoc Jr, Z. W. Lim, and R. Ling, Defining fake news A typology of scholarly definitions, Digital journalism, 6(2):137-153, 2018.
[14] X. Zhou, A. Jain, V. V. Phoha, and R. Zafarani, 'Fake news early detection: A theory-driven model', Digital Threats: Research and Practice, vol. 1, no. 2, pp. 1–25, 2020.
[15] X. Zhou and R. Zafarani, 'A survey of fake news: Fundamental theories, detection methods, and opportunities', ACM Computing Surveys (CSUR), vol. 53, no. 5, pp. 1–40, 2020.
[16] K. Shu, S. Wang, and H. Liu, Beyond news contents: The role of social context for fake news detection, Proceedings of the 12th ACM International Conference on Web Search and Data Mining, 312–320,2019.
[17] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, 'BERT: Pre-training of deep bidirectional transformers for language understanding', arXiv preprint arXiv:1810. 04805, 2018.
[18] J. W. Pennebaker, M. E. Francis, and R. J. Booth, 'Linguistic inquiry and word count: LIWC 2001', Mahway: Lawrence Erlbaum Associates, vol. 71, no. 2001, p. 2001, 2001.
[19] M. Szczepański, M. Pawlicki, R. Kozik, and M. Choraś, 'New explainability method for BERT-based model in fake news detection', Scientific Reports, vol. 11, no. 1, pp. 1–13, 2021.
[20] 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, 80(8): 11765–11788, 2021.