Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 239 (2024) 2056–2064

Procedia Computer Science

www.elsevier.com/locate/procedia

CENTERIS – International Conference on ENTERprise Information Systems / ProjMAN – International Conference on Project MANagement / HCist – International Conference on Health and Social Care Information Systems and Technologies 2023

Does Fake News have Feelings?

Herbert Laroca[a,b,*], Vitor Rocio[b], António Cunha[a,c]

[a]*University of Trás-Os-Montes e Alto Douro, Vila Real, Portugal*
[b]*Universidade Aberta, Lisboa, 1250-100, Portugal*
[c]*INESC Technology and Science (INESC-TEC), Porto, Portugal*

Abstract

Fake news spreads rapidly, creating issues and making detection harder. The purpose of this study is to determine if fake news contains sentiment polarity (positive or negative), identify the polarity of sentiment present in their textual content and determine whether sentiment polarity is a reliable indication of fake news. For this, we use a deep learning model called BERT (Bidirectional Encoder Representations from Transformers), trained on a sentiment polarity dataset to classify the polarity of sentiments from a dataset of true and fake news. The findings show that sentiment polarity is not a reliable single feature for recognizing false news correctly and must be combined with other parameters to improve classification accuracy.

Keywords: fake news;sentiment analysis;deep learning;BERT.

1. Introduction

Today's networked world benefits society and numerous opportunities have arisen, provided by the creation of the internet and its technologies. Therefore, information gains a central role in the information society, becoming a

* Corresponding author.
E-mail address: herbert.laroca@gmail.com

10.1016/j.procs.2024.06.392

valuable and disputed good. Despite the benefits, the network society also brings several challenges, including the negligent use of information for illicit purposes. Several works [21,2] were developed to deal with false information and its effects. The information disorder, as cited by Tandoc Jr et al. [21], is a relevant and urgent topic today, representing a set of actions and effects caused by the dissemination of false information. Information disorder can generally be categorized into disinformation, misinformation, and malinformation [21]. Disinformation is false information spread intentionally and with a defined purpose. Misinformation is false information shared unintentionally and inadvertently. Malinformation is truthful information shared with the intent to cause harm. False information and information disorder are generally linked to the buzzword fake news [21]. The term was defined by news media and quickly spread like its own content.

Despite being initially linked to journalism and political statements, today, fake news affects various areas of society, causing great financial damage [6]. As cites Gelfert [10, p.85],” Fake news is the deliberate presentation of (typically) false or misleading claims as news, where the claims are misleading by design”. Fake news is categorized as a type of disinformation. According to Liu et al. [16], the textual information available in the media, including the Internet, can be categorized between facts and opinions. Liu et al. [16, p.1] show that ” facts are objective expressions about entities, events, and their properties. Opinions are generally subjective expressions that describe people’s feelings, evaluations, or feelings toward entities, events, and their properties”. Before the emergence of the web, factual content generated in the text was much larger than opinionated content but with the appearance of social networks and online review sites, this situation has been balanced [16].

According to Liu [15, p.7],” sentiment analysis and opinion mining is the field of study that analyses people’s opinions, sentiments, evaluations, attitudes, and emotions from written language”. Analyzing the sentiment of a given text, apparently opinionated, is to find out if its written content expresses a positive, neutral, or negative sentiment about a certain subject. The work of Bowman and Cohen [5] points out a relationship between fake news and emotions by showing that news with strong emotional appeals tends to be remembered and spread. The authors deduce that false news and information have an emotional charge, making them more likely to be shared.

A large part of the success of fake news lies in its ability to draw the collective’s attention from its emotions. This work aims to verify if texts considered fake news contain sentiment polarity (positive or negative), defining a way of classifying fake news based on sentiment analysis.

This paper is organized into the following sections: The related work section presents similar works that use sentiment analysis to detect fake news and present its most important characteristics. The research methodology section presents the scientific methodological aspects of the paper and the research questions. The practical section methodology lists the praxis steps performed during the research development, with a description of the datasets, infrastructure, and model. The discussion section presents an evaluation and discussion of the results found, ending with the conclusion section, where the research questions are answered and future research is demonstrated.

2. Related Work

Fake news classification proposals can be categorized into three approaches: (i) knowledge-based approaches and their relationships, using knowledge bases; (ii) context-based approaches, which use metadata about the information to verify its veracity, such as authorship, publisher, labeled target audience, comments, etc.; (iii) approaches based on the own content of the news, represented mainly by text [13]. Sentiment analysis, the focus of this work, is an approach based on informational content represented by textual expressions of sentiment that may be included in the text. The related works evaluated are listed in Table 1. In general, the works that associate Fake News with sentiment analysis use two types of strategy: use sentiment analysis as the main feature to classify fake news [14, 19, 7, 9, 18, 23], and use of sentiment analysis as additional support for classifying fake news, evaluating sentiment as an additional feature [4, 22, 20, 13].

Most of the works evaluated use methods, techniques, and technologies related to natural language processing, like Bi-grams, Tri-grams, TF-IDF [4], Tokenization [14, 20], Topic Modelling [22, 9], Stemming [20], and traditional machine learning algorithms like Naïve Bayes [4, 9, 18], Random Forest [4, 18, 20], Support Vector Machine (SVM) [22, 18, 20], K-Nearest Neighbors (KNN) [22, 20], Decision Trees [22, 9, 20] and Logistic Regression [22, 20]. Some works use neural networks models to classify the data like Deep Convolutional Neural Networks (CNN) [7], Bi-Directional Long Short-Term Memory (Bi-LSTM) [9], Recurrent Neural Networks (RNN) [13], Gated Recurrent

Units (GRU) [13] and Long Short-Term Memory (LSTM) [13]. Several works compare classification algorithms [4, 9, 18, 20, 13) and some use specific tools for sentiment analysis, like IBM Watson API [19], Dandelion API [22], MeaningCloud TextBlob, and AFINN [23]. Regarding data, the evaluated works use datasets provided by the Kaggle platform [4, 14, 18] and by sites such as PolitiFact [4, 7, 9, 23], Snopes [19] and GossipCop [7, 9, 23].

Table 1. - Related Works

Title	Year	Author	T1	T2	T3
Fake News Detection Using Sentiment Analysis	2019	[4]	1	1	2
Fake News Identification Based on Sentiment and Frequency Analysis	2019	[14]	3	1,8	1
Investigating the emotional appeal of fake news using artificial intelligence and human contributions	2019	[19]	2	4,5	1
Polarization and Fake News: Early Warning of Potential Misinformation Targets	2019	[22]	2	7,2	2
SAME: Sentiment-aware multi-modal embedding for detecting Fake news	2019	[7]	1	4,3	1
Fake News Classification of Social Media Through Sentiment Analysis	2020	[9]	3	1,7,2,3	1
Comparison of Machine Learning Algorithms for Sentiment Classification on Fake News Detection	2021	[18]	2	2	1
On Sentiment of Online Fake News	2021	[23]	1	4	1
Sentiment Analysis for Fake News Detection	2021	[3]	2	6	3
Sentiment Analysis Model for Fake News Identification in Arabic Tweets	2021	[20]	1	1,2	2
Covid-19 fake news sentiment analysis	2022	[13]	2	3	2

(T1: 1 = Conference Paper 2 = Journal Article 3 = Book Section / T2: 1 = NLP 2 = Machine Learning Models 3 = Deep Learning Models 4 = Sentiment Analysis / Tools 5 = Human in the Loop 6 = Review 7 = Topic Modelling 8 / T3: 1 = Primary Feature 2 = Secondary Feature 3 = Primary and Secondary features)

Regarding the results obtained and the relationship between the sentiment and its polarity (negative, neutral, or negative), the evaluated works converge to negative polarity in fake news. The work of Kapusta et al. [14] shows that fake news articles had a more negative sentiment. Paschen [19] mentions that fake news titles are more emotional and less neutral about feelings than true news titles, also revealing that negative sentiments are stronger in fake news than true news, both in titles and main texts. Ding et al. [9] go further, showing that the textual content of the tweets analyzed in their work tends to contain more sentiment than neutrality, with a tendency towards positive sentiment for true news and negative sentiment for fake news. The authors also claim that the texts' subject can influence the patterns of sentiment to be verified in the news, both true and false. Alonso et al. [3] also confirm the same relationship, indicating that fake news has a negative sentiment and real news has a positive sentiment. Most of the evaluated works make the classification of sentiments from the evaluated texts themselves, that is, the polarity of sentiment is extracted from the texts considered fake news. This work makes a different approach, using the BERT neural network to train a model using a dataset already prepared for sentiment analysis. Only after training the model with the sentiment dataset is it used to classify the sentiment polarity of the texts from a fake news dataset. The proposed experiment aims to apply a neural network model based on BERT trained in a sentiment dataset to classify sentiment polarity in a dataset of false and true news, allowing the analysis between sentiment polarity and fake news.

3. Methodology

3.1. Research Questions.

This work proposes the following research questions (RQ1): Can fake news be identified from the polarity of sentiment existing in its textual content? (RQ2): Is sentiment polarity an indicator of fake news?

3.2. Research Methodology.

This work presents characteristics of applied research with the use of quantitative and exploratory statistical methods. From the collection of data available in two public datasets, which are detailed in section 3.3, a controlled experiment is made to extract quantitative statistical data to obtain information that allows answering the proposed research questions. The practical development of the research follows the steps defined and illustrated in Figure 1, detailed in the following sections. After collect the input data responsible for the supervised training of the neural network, the pre-processing step is made to adapt the formats required by the neural network. After data process, the neural network training and classification process is performed. During the classification step the data from a fake news dataset are used to do the classification. Then the classified data are turned available for visualization and analysis. To allow the execution of the experiment, the infrastructure of Google Cloud virtual machines was used. The configuration of the virtual machine used was: Intel(R) Xeon(R) CPU @ 2.30GHz with 50GiB storage and 7680 MiB ram, Tesla k80 GPU and Debian Operating System. Furthermore, Keras, Tensorflow, and Jupyter Notebook software were used to develop applications under Python.

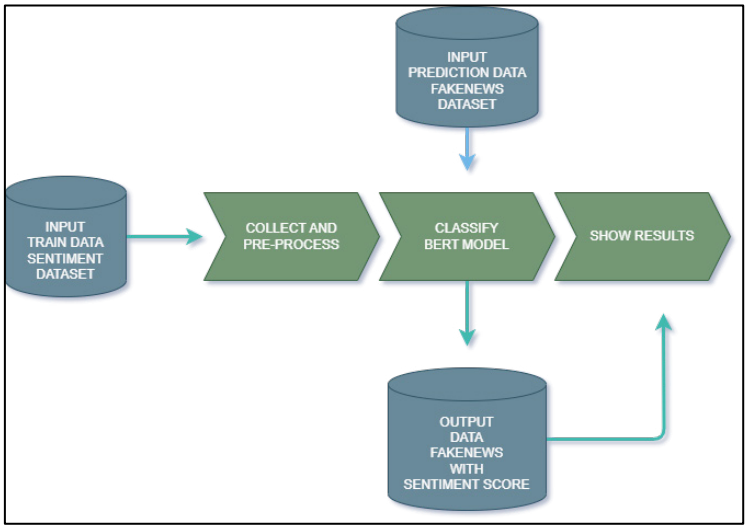


Fig. 1. Graphical representation from the research steps (authors).

The input data used in the research are divided into train and predict data, which are the data that the neural network will classify. The training dataset contains negative and positive sentiment polarity and the dataset to be classified includes true and fake news. The dataset named IMDB-reviews [17] was used to train the model. This dataset contains 25,000 movie reviews for training and 25,000 reviews for testing, split between negative and positive sentiment polarity. The dataset is available directly from the TensorFlow datasets collection [11], being collected by its function (load-dataset). An example of the reviews is shown in Table 2. The data for classifying false and true news was obtained from the ISOT dataset [1,2,12]. The real news articles were obtained from the news site Reuters.com. Fake news articles were collected from different sources and flagged by PolitiFact and Wikipedia.

Table 2. Example of train file texts from IMDB reviews dataset [17]

Text	Polarity
Bromwell High is a cartoon comedy. It ran simultaneously with other programs about school life, such as "Teachers". My 35 years in the teaching profession led me to believe that bromwell High's satire is much closer to reality than "Teachers". The scramble to survive financially, the insightful students who can see right through their pathetic teachers' pomp, and the pettiness of the whole situation all remind me of the schools I knew and their students. When I saw the episode in which a student repeatedly tried to burn down the school, I immediately recalled ..... at ..... High. A classic line: INSPECTOR:	Pos

I'm here to sack one of your teachers. STUDENT: Welcome to Bromwell High. I expect many adults my age to think Bromwell High is far-fetched. What a pity that it isn't!

The dataset contains various articles on various topics. However, most articles focus on political and world news. The first file, called True.csv, has more than 21000 articles and the second file, called Fake.csv, contains more than 23000 articles. Table 3 shows the columns of the ISOT dataset used in the research to classify fake news based on the polarity of sentiment. The table 4 shows an example of textual data contained in the fake.csv file of the ISOT dataset. As already mentioned, data collection for training was performed directly by the load dataset functionality of the TensorFlow tool. The fake news classification dataset was collected by reading files in CSV (comma-separated values) format.

Table 3. The columns of the fake news dataset [1].

Column	Description
title	The title of the article
text	The text of the article
subject	The subject of the article

Before the model training stage, performing some filtering and simplification tasks in the texts was necessary to allow better use and data processing optimization. This stage, called pre-processing, has the following tasks: (i) normalize text to lowercase; (ii) remove the references to URLs (Uniform Resource Locators) in the texts of the reviews; (iii) remove the numbers; (iv) remove punctuation and (v) remove words and numbers with less than three characters. Text pre-processing is an important task to avoid errors and overloads during dataset processing and training. Table 5 shows an example of the text after the processing tasks. The data from the sentiment polarity dataset is provided for training the neural network.

Table 4. Examples of fake news file texts [1].

Title	Text	Subj	Date
Donald Trump Sends Out Embarrassing New Year's Eve Message;This is Disturbing	Donald Trump could not wish all Americans a Happy New Year and leave it at that. Instead, he had...	News	2017
Drunk Bragging Trump Staffer Started Russian Collusion Investigation	House Intelligence Committee Chairman Devin Nunes will have a bad day. He s been under the as...	News	2017

After training, the model is used to classify sentiment polarity in the fake news dataset. Sentiment polarity is represented by a score from 0 to 1, indicating that negative sentiment is expressed by scores close to 0 and positive sentiment is represented by scores close to 1.

Table 5. Example of text processed from IMDB datasets folder [17]

Original Text	Processed Text
Bromwell High is a cartoon comedy. It ran simultaneously with other programs about school life, such as "Teachers". My 35 years in the teaching profession led me to believe that Bromwell High's satire is much closer to reality than "Teachers". The scramble to survive financially, the insightful students who can see right through their pathetic teachers' pomp, the pettiness of the whole situation.	Bromwell high cartoon comedy same time, some other programs about school life, such as teachers years teaching profession lead believe that Bromwell highs satire much closer reality than teachers scramble survive financially insightful students right through their pathetic teacher's pomp pettiness whole situation

The machine learning model used in this work is called BERT [8]. BERT stands for Bidirectional Encoder Representations from Transformers. This model was proposed by Jacob Devlin and his Google research colleagues in 2019 [8] and is very similar to a transformer, more specifically, like a transformer encoder. Nowadays, the BERT model has reached the state of the art for some tasks in NLP (Natural Language Processing). The advantages of BERT

lie in (i) pre-training the linguistic model on a large volume of unlabeled data, allowing a simple fine-tuning without drastically modifying its architecture, which is unique and performs many tasks (figure 2); and (ii) the use of two-word prediction tasks during the pre-training of the model - Masked LM: This task randomly and bidirectionally masks some of the tokens from the input to predict the masked word based only on its context - Next Sentence Prediction (NSP): which jointly pre-trains representations of text pairs [8]. For instance, BERT has shown to be an excellent solution for problems related to sentiment analysis. The model used has a BERT layer pre-trained using data from Wikipedia and is available in the TensorFlow hub.

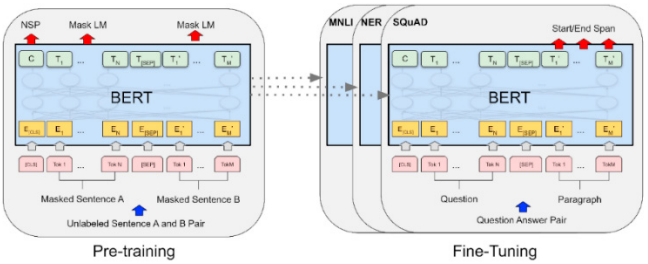


Fig. 2. Overall pre-training and fine-tuning procedures for BERT [8].

4. Results

After classifying the news, numbers between 0 and 1 were generated, representing the sentiment polarity scores for twenty thousand fake news and true news and their respective titles. The scores were separated into a fake news file and a real news file. The polarity was set to 0 for negative and 1 for positive sentiment. Scores less than 0.5 were classified as negative polarity, and scores above 0.5 were classified as positive polarity. An example of the scores and polarities created in each file is illustrated in Table 6.

Table 6. Example of scores and polarities created (Authors)

Score	Pol
0.0477975569665432	0
0.6508778929710388	1

Table 7 shows the average and median of true and false news scores, separated by title and text. The average is the sum of all values in a data set divided by the number of elements in the same set and the median represents the central value of an ordered set of data.

Table 7. Average and median of scores (Authors)

News Type	**Average Score**	**Median Score**
Real News (title)	0.68	0.52
Real News (text)	0.52	0.83
Fake News (title)	0.52	0.29
Fake News (text)	0.43	0.59

Table 8 presents the precision and recall measures obtained in the data set. Precision (1) was obtained by calculating the number of fake news with negative polarity divided by the amount of news with negative polarity. The recall (2) was obtained by calculating the number of fake news with negative polarity divided by the number of fake news.

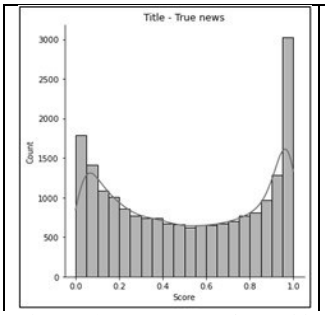
Table 8. Fake News Precision and Recall measures (Authors).

Fake News	Precision	Recall
Title	0.54	0.59
Text	0.62	0.48

$$precision = \frac{number\ of\ fake\ news\ with\ negative\ polarity}{number\ of\ news\ with\ negative\ polarity} \quad (1)$$

$$recall = \frac{number\ of\ fake\ news\ with\ negative\ polarity}{number\ of\ fake\ news} \quad (2)$$

The figures below represent the distribution of scores along with true news (figure 3a and figure 3b) and fake news (figure 4a and figure 4b). In the next sections, the results will be analyzed and discussed.



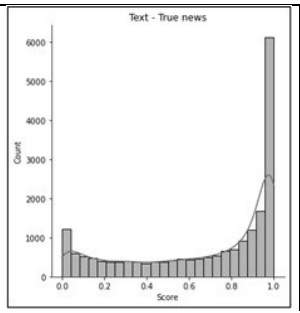
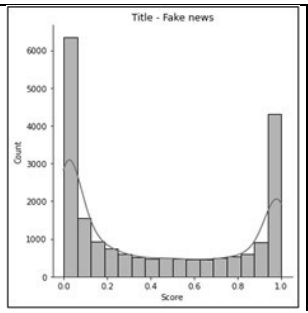


Fig. 3. (a) True news and quantities from the title; (b) True news and quantities from text.



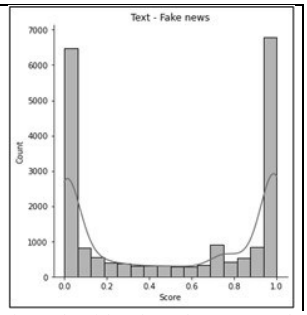


Fig. 4. (a) Fake news and quantities from the title; (b) Fake news and quantities from text.

5. Discussion

According to the related work (section 2), there is a tendency for a negative classification of fake news and a positive classification of true news. The average and median scores, shown in Table 7, were calculated to identify a general trend in polarity in each data set. It is observed by the values that the texts of fake news have an average score of 0.43, and the titles of fake news have an average score of 0.52. It indicates that they have slightly negative polarity. In contrast, the values of the true news texts have an average score of 0.52, and the true news titles have an average score of 0.68, indicating a positive polarity. The medians follow the same trend, showing high positive median scores for news (0.83), slightly positive (0.52) true titles with slightly negative median scores for news (0.59) and low for titles (0.29).

According to the values found for accuracy, 54% of news titles classified with negative polarity are fake news, and 62% of news texts classified with negative polarity are fake news. Regarding the recall, 59% of fake news titles were classified with negative polarity and 48% of fake news texts were classified with negative polarity. It is noted on the histogram presented in Figure 3a and Figure 3b that the frequency is higher in scores close to one, indicating a trend towards positive polarity, both for titles (figure 3a) and texts (figure 3b). Concerning the histogram of the scores and their distribution in fake news (figure 4a and figure 4b), there is more fake news with the sentiment, regardless of whether they are positive or negative. Fake news seems to have sentiments in its text and titles. The data indicate that, in general, fake news tends to be sentimental, regardless of its polarity. Previous studies have found that fake news has a negative polarity, while real ones have a positive polarity. If we look at the graphics and averages of true news, there is a general trend toward positive polarity. The trend is not confirmed in the graphics and averages of fake news. The graphics show that fake news has a polarity of sentiment, both positive and negative.

6. Conclusion

The objective of this introductory study was to verify if there is a relationship between the polarity of sentiment and fake news to support the process of combatting the dissemination of false information. To experiment, a BERT deep learning neural network model was trained to classify a group of fake and true news. The neural network was trained with a sentiment dataset, and the classification was performed on the fake and true news datasets.

After the experiment, we can answer the research questions: (RQ1) Is it possible to identify fake news from the polarity of sentiment existing in its textual content? According to our experiment, it is not possible to identify fake news with precision from the polarity of sentiment existing in its textual content. The polarity of sentiment, both positive and negative, may suggest the existence of fake news and not fact. Therefore, sentiment polarity should not be considered only a unique good indicator of fake news. The use of sentiment analysis as a feature must be evaluated in conjunction with other features to increase the classification accuracy. (RQ2) Is sentiment polarity an indicator of fake news? The sentiment is indeed present in fake news, although it has shown to be a not-so-good unique indicator. Because most of the sentiment analysis datasets are based on reviews, our work used an already-known base of movie reviews, limiting and influencing the results. For future work, we intend to use a specific news dataset to check whether the scores are maintained or changed. It is also intended to classify news such as politics, varieties, science, etc. to compare scores. Another investigative opportunity lies in verifying factual or opinionated content.

Acknowledgements

This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia, within project LA/P/0063/2020.

References

[1] Ahmed, H., Traore, I., Saad, S. (2017). Detection of online fake news using n-gram analysis and machine learning techniques, in: International conference on intelligent, secure, and dependable systems in distributed and cloud environments, Springer. pp. 127–138.

[2] Ahmed, H., Traore, I., Saad, S. (2018a). Detecting opinion spams and fake news using text classification. Security and Privacy 1, e9.

[3] Alonso, M.A., Vilares, D., Gomez-Rodrıguez, C., Vilares, J. (2021). Sentiment analysis for fake news detection. Electronics 10, 1348.

[4] Bhutani, B., Rastogi, N., Sehgal, P., Purwar, A. (2019). Fake news detection using sentiment analysis, in: 2019 twelfth international conference on contemporary computing (IC3), IEEE. pp. 1–5.

[5] Bowman, N.D., Cohen, E. (2020). Mental shortcuts, emotion, and social rewards: the challenges of detecting and resisting fake news. Fake news: Understanding media and misinformation in the digital age , 223–233.

[6] Cambria, E., Das, D., Bandyopadhyay, S., Feraco, A. (2017). Affective computing and sentiment analysis, in: A practical guide to sentiment analysis. Springer, pp. 1–10.

[7] Cui, L., Wang, S., Lee, D. (2019). Same: sentiment-aware multi-modal embedding for detecting fake news, in: Proceedings of the 2019 IEEE/ACM international conference on advances in social networks analysis and mining, pp. 41–48.

[8] Devlin, J., Chang, M.W., Lee, K., Toutanova, K. (2019). BERT: Pretraining of deep bidirectional transformers for language understanding, in: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), Association for Computational Linguistics, Minneapolis, Minnesota. pp. 4171–4186. URL: https://aclanthology.org/N19-1423, doi:10.18653/v1/N19-1423.

[9] Ding, L., Ding, L., Sinnott, R.O. (2020). Fake news classification of social media through sentiment analysis, in: International Conference on Big Data, Springer. pp. 52–67.

[10] Gelfert, A. (2018). Fake news: A definition. Informal logic 38, 84–117.

[11] Google (2022). Tensorflow datasets catalog - imdb reviews 2022.

[12] ISOT Research Lab (2018). Isot fake news dataset.

[13] Iwendi, C., Mohan, S., Ibeke, E., Ahmadian, A., Ciano, T., et al. (2022). Covid-19 fake news sentiment analysis. Computers and electrical engineering 101, 107967.

[14] Kapusta, J., Benko, L., Munk, M. (2019). Fake news identification based on sentiment and frequency analysis, in: International Conference Europe Middle East & North Africa Information Systems and Technologies to Support Learning, Springer. pp. 400–409.

[15] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis lectures on human language technologies 5, 1–167.

[16] Liu, B., et al. (2010). Sentiment analysis and subjectivity. Handbook of natural language processing 2, 627–666.

[17] Maas, A., Daly, R.E., Pham, P.T., Huang, D., Ng, A.Y., Potts, C. (2011). Learning word vectors for sentiment analysis, in: Proceedings of the 49th annual meeting of the association for computational linguistics: Human language technologies, pp. 142–150.

[18] Mahmud, Y., Shaeeali, N.S., Mutalib, S. (2021). Comparison of machine learning algorithms for sentiment classification on fake news detection. International Journal of Advanced Computer Science and Applications 12.

[19] Paschen, J. (2019). Investigating the emotional appeal of fake news using artificial intelligence and human contributions. Journal of Product & Brand Management.

[20] Sawan, A., Thaher, T., et al. (2021). Sentiment analysis model for fake news identification in arabic tweets, in: 2021 IEEE 15th International Conference on Application of Information and Communication Technologies (AICT), IEEE. pp. 1–6.

[21] Tandoc Jr, E.C., Lim, Z.W., Ling, R. (2018). Defining "fake news" a typology of scholarly definitions. Digital journalism 6, 137–153.

[22] Vicario, M.D., Quattrociocchi, W., Scala, A., Zollo, F. (2019). Polarization and fake news: Early warning of potential misinformation targets. ACM Transactions on the Web (TWEB) 13, 1–22.

[23] Zaeem, R.N., Li, C., Barber, K.S. (2020). On sentiment of online fake news, in: 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), IEEE. pp. 760–767.