

CERIST Natural Language Processing Challenge

March 29th, 2023

Modeling Fake News Detection Using Machine Learning Algorithms for Arabic covid-19 Tweets

Mohammed Aldawsari^a, Omer Salih Dawood Omer^a, Yousra F.G. Elhakeem^b, Safa Eltayeb^{c*}

^a*Department of Computer Science, College of Arts and Science in Wadi Aldawasir, Prince Sattam bin Abdulaziz University, Wadi-Aldawasir, Saudi Arabia*

^b*Yanbu, Saudi Arabia*

^c*Department of Software Engineering, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, P.O. Box 151, Alkharj 11942, Saudi Arabia*

Abstract

Fake news detection has become a major issue in the digital age, with social media playing a major role in its spread. This paper outlines the dataset and methodology used to model Arabic fake news. This paper is about our participation on *CERIST Natural Language Processing Challenge*. We used the dataset provided for the Task1.c. Arabic sentiment analysis and fake news detection within covid-19. The model used for this task is a simple transformer fake news model based on the Arabic pre-trained language model CAMEL-BERT. This model was utilized in two variants: a fine-tuned model and a Bidirectional long short-term model. The experiment results of this modeling CAMEL-BERT provides the best result by achieving 0.959 F1, thus outperforming all other models variants in detecting fake news.

Keywords: Fake News Detection; Machine Learning; CAMEL-BERT; Covid-19; Twitter; CERIST.

* E-mail address: Dr.yousraelhakeem@gmail.com.

1. Introduction

Nowadays social media plays an important role as source of news due to (i) more timely and less expensive to consume compared with traditional news media, and (ii) it is easier to share, comment on, and discuss news with friends or other social media, Shu et al., 2017. Many nations depend on social media as major source of news, 62 percent of U.S. adults get news on social media in 2016, Shu et al., 2017. Even in developing countries about 62 percent of internet users depend on social media for getting news Elhadad et al., 2019. Fake/False defined news as “news articles that are potentially or intentionally misleading for the readers, as they are verifiable and deliberately” Pierri and Ceri, 2019. As shown in Figure 1, research directions of fake news can be categorized in four categories: Data oriented, Feature-oriented, Model-oriented and Application oriented Shu et al., 2017.

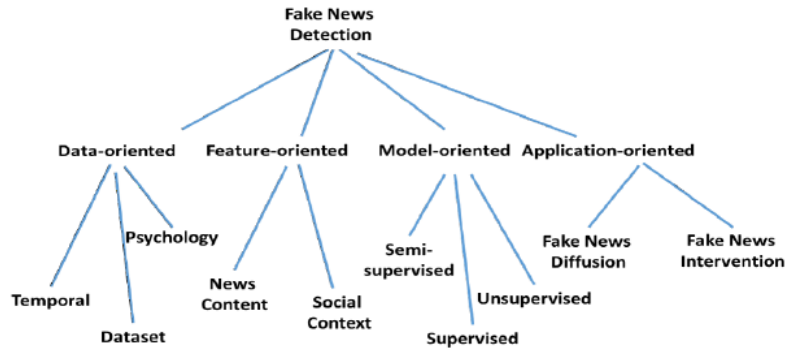


Fig. 1. Future directions and open issues for fake news detection on social media (Shu et al., 2017).

2. Related Work

There are many researches, tools, and methodologies covered area of fake news detection. A tool was designed with the specific aim of detecting and eliminating web pages that contain misinformation intended to mislead reader, tool installed in web browser detect and filter out potential Click-baits Aldwairi and Alwahedi, 2018. A simple approach for fake news detection with the help of K- Nearest Neighbor, algorithm predict fake news or real news on basis of learning behavior, Kesarwani et al., 2020. Unsupervised fake news detection on social media was investigated by exploiting the users' unreliable social engagement information, and an unsupervised learning framework, UFD, was proposed, which utilizes a probabilistic graphical model to model the truths of news and the users' credibility, Yang et al., 2019. A fake news detection model was proposed to detect fake news in early stages, it consists of four main components, propagation path construction and transformation, RNN-based propagation path representation, CNN-based propagation path representation, and propagation path classification, which are integrated together to detect fake news at the early stage of its propagation Liu and Wu, 2018. The authors in Shu et al., 2017 present a comprehensive overview of the methods used to analyze and detect fake news on Twitter. The authors discuss the various techniques used to identify trends in information, such as sentiment analysis, topic modeling, and network analysis. They also discuss methods for detecting fake news, including machine learning algorithms, natural language processing techniques, and content-based approaches. Finally, the authors provide an overview of existing tools for analyzing and detecting fake news on Twitter. The paper provides a valuable resource for researchers interested in understanding how to effectively analyze and detect fake news on Twitter. Lee and Chua, 2022, presents a novel approach to automatically detect fake news on Twitter by combining linguistic

and psycholinguistic features with machine learning algorithms. The authors propose a feature set that includes lexical, syntactic, semantic, and psychological features. They also present an experimental evaluation of their proposed method on two datasets of tweets related to the 2016 US presidential election. The results show that the proposed method outperforms existing methods in terms of accuracy, precision, recall, and F1-score. The paper concludes by discussing the implications of the findings for future research in this area. The authors in Zervopoulos et al., 2022 used a convolutional neural network (CNN) to classify tweets as either real or fake news. The model was trained on a dataset of over 10,000 tweets collected from Twitter during the protests. The results showed that the CNN model achieved an accuracy of 97.3%, precision of 96.8%, recall of 97.7%, and F1 score of 97.2%. This demonstrates the effectiveness of deep learning for detecting fake news on Twitter related to the 2019 Hong Kong protests. In Patra et al., 2022, the authors provide a comprehensive survey of the current state of research on fake news detection on social media. It covers a range of topics, including the various methodologies used to detect fake news, the datasets used to train and evaluate these methods, and the challenges faced in this field. The paper also discusses some of the potential applications of fake news detection, such as improving public discourse and increasing trust in online sources.

3. Experiments

This section outlines the dataset and methodology that has been utilized for modeling Arabic fake news. For the dataset, we use the shared Arabic fake news dataset Hadj Ameur and Aliane, 2021, and randomly split this dataset into 75% train, and 25% validation sets. For the methodology, the model we use for this task is a simple transformer fake news model based on the Arabic pre-trained language model CAMEL-BERT (Inoue et al., 2021). We utilize this model in two variants that are fine-tuned model and Bidirectional long short term model (BiLSTM). For the first, we fine-tuned the CAMEL-BERT model on three parameters batch size, learning rate, and number of epochs as shown in table 1. The obtained results from this model is shown in table 2. For the second experiment, we extract embeddings from the CAMEL-BERT and trained a BiLSTM model with the three aforementioned parameters (i.e., batch size, learning rate, and number of epochs). The obtained results from this model is shown in table 2.

We also utilized another pre-trained language model the so-called Bert-base-multilingual-cased (Devlin et al., 2018) with the same methodologies for comparison. The result of using this model is shown on table 2. From the results in table 2, it is clear that Fine-Tuning CAMEL-BERT provides the best result by achieving 0.959 F1, thus outperforming all other models variants in detecting fake news.

Table 1. The parameters used in all experiments

Model	Method	Batch size	Learning rate	Number of epochs
CAMEL-BERT	Fine-Tune	32	0.0001	2
	BiLSTM	64	0.001	10
Bert-base-multilingual-cased	Fine-Tune	32	0.0001	2
	BiLSTM	64	0.001	10

Table 2. Models Results

Model	Method	F1
CAMEL-BERT	Fine-Tune	0.959
	BiLSTM	0.923

Bert-base-multilingual-cased	Fine-Tune	0.92
	BiLSTM	0.90

4. Conclusion

Fake news detection using machine learning algorithms in tweets is important because it can help to identify and combat the spread of false information on social media. Social media platforms have become a major source of news and information, and the spread of false information can have serious consequences. Machine learning algorithms can help to detect fake news by analyzing the content of tweets, such as the language used, sentiment analysis, and other features. This can help to identify tweets that contain false information or are intended to deceive or mislead users. In this paper, we have demonstrated that fine-tuning the CAMEL-BERT model is an effective approach for detecting fake news. The results obtained from this model outperformed all other models variants in detecting fake news. We also showed that using a pre-trained language model such as Bert-base-multilingual-cased can provide good results as well. This research provides a valuable insight into the potential of utilizing pre-trained language models for fake news detection.

References

- Aldwairi, M., & Alwahedi, A., 2018. Detecting fake news in social media networks. *Procedia Computer Science*, 141, 215-222.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Elhadad, M. K., Li, K. F., & Gebali, F., 2019. Fake news detection on social media: a systematic survey. In *2019 IEEE Pacific Rim conference on communications, computers and signal processing (PACRIM)* (pp. 1-8). IEEE.
- Hadj Ameur M.S., Aliane H., 2021. AraCOVID19-MFH: Arabic COVID-19 multi-label fake news & hate speech detection dataset. *Procedia Computer Science* 189, 232-241.
- Inoue, G., Alhafni, B., Baimukan, N., Bouamor, H., & Habash, N., 2021. The interplay of variant, size, and task type in Arabic pre-trained language models. *arXiv preprint arXiv:2103.06678*.
- Kesarwani, A., Chauhan, S. S., & Nair, A. R., 2020. Fake news detection on social media using k-nearest neighbor classifier. In *2020 international conference on advances in computing and communication engineering (ICACCE)* (pp. 1-4). IEEE.
- Lee, C. J., & Chua, H. N., 2022. Using Linguistics and Psycholinguistics Features in Machine Learning for Fake News Classification Through Twitter. In *Proceedings of International Conference on Data Science and Applications: ICDSA 2021, Volume 1* (pp. 717-730). Springer Singapore.
- Liu, Y., & Wu, Y. F., 2018. Early detection of fake news on social media through propagation path classification with recurrent and convolutional networks. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 32, No. 1).
- Patra, D., Jana, B., Mandal, S., & Sekh, A. A., 2022. Understanding Fake News Detection on Social Media: A Survey on Methodologies and Datasets. In *International Symposium on Artificial Intelligence* (pp. 226-242). Cham: Springer Nature Switzerland.
- Pierri, F., & Ceri, S., 2019. False news on social media: a data-driven survey. *ACM Sigmod Record*, 48(2), 18-27.
- Yang, S., Shu, K., Wang, S., Gu, R., Wu, F., & Liu, H., 2019. Unsupervised fake news detection on social media: A generative approach. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 33, No. 01, pp. 5644-5651).
- Shu, K., Sliva, A., Wang, S., Tang, J., & Liu, H., 2017. Fake news detection on social media: A data mining perspective. *ACM SIGKDD explorations newsletter*, 19(1), 22-36.
- Zervopoulos, A., Alvanou, A. G., Bezas, K., Papamichail, A., Maragoudakis, M., & Kermanidis, K., 2022. Deep learning for fake news detection on Twitter regarding the 2019 Hong Kong protests. *Neural Computing and Applications*, 34(2), 969-982.