1 Literature Review

Much of the previous research Capuano et al. (2023) on fake news detection has been exploratory with neural networks on labeled data. Ma et al. (2016) first introduced the deep learning methods (RNN, LSTM and GRU) to the fake news detection domain in 2016. This paper introduces an innovative method leveraging recurrent neural networks (RNN) to autonomously identify rumors. The proposed model utilizes RNNs to learn continuous representations of microblog events, capturing the evolving contextual information over time. Results from experiments on datasets from real-world microblog platforms reveal that the RNN-based method surpasses existing rumor detection models reliant on hand-crafted features. Additionally, enhancements in performance are achieved through advanced recurrent units and extra hidden layers. The RNN-based approach demonstrates quicker and more accurate rumor detection compared to established techniques, including leading online rumor debunking services. The precision for this model is 94.7% and recall is 96.3% on Twitter dataset.

What's more, machine learning approaches are also popular methods in dealing with Fake news detection (Ahmed et al., 2021). Manzoor et al. (2019) conducts an in-depth analysis of existing research on fake news detection, focusing on traditional machine learning models. The objective is to select the most effective model for developing a product that utilizes supervised machine learning algorithms to classify news articles as true or false. The proposed approach involves leveraging tools such as Python's scikit-learn and natural language processing (NLP) for textual analysis. The process encompasses feature extraction and vectorization, employing scikit-learn's Count Vectorizer and Tiff Vectorizer for tokenization and feature extraction. Additionally, feature selection methods are employed to identify the most suitable features that yield optimal precision, guided by the results of a confusion matrix.

Yu et al. (2017) proposes a novel method, the Convolutional Approach for Misinformation Identification (CAMI) based on the Convolutional Neural Network (CNN). CAMI can flexibly extract key features scattered among an input sequence and shape high-level interactions among significant features, which help effectively identify misinformation and achieve practical early detection. Experiment results on two large-scale datasets validate the CAMI model's effectiveness on misinformation identification and early detection tasks, with an accuracy of 93.3% and 77.7% respectively on the two datasets.

Many of other work have been successfully built up based on supervised perspective (Ma et al., 2019; Vaibhav et al., 2019). And there are many other approaches using weakly-supervised methods. Konkobo et al. (2020) trained a supervised model and an unsupervised model. It utilizes not only the news content information but also the user's comment information on the news and the author's credibility information. Mansouri et al. (2020) used the LDA method to pseudo-label unlabeled data for better training of unlabeled CNN models. The precision of this method is 95.6%, and the recall is 96.7%.

2 Dataset description

In our project, we are aiming to discover two main datasets and to find out the model performance of combining them.

The first one is FNC-1 dataset [1]. It is a well-known dataset in natural language processing

[1] http://www.fakenewschallenge.org/

and machine learning. The Fake News Challenge is an initiative that aims to explore and improve the automatic detection of fake news. The FNC-1 dataset specifically focuses on the task of stance detection, where the goal is to determine the stance of a body of text (typically a news article) with respect to a headline. The dataset was released as part of a competition.

The second one we are using is LIAR dataset [2]. It is a dataset created for the purpose of advancing research in the field of fake news detection. It was introduced in the paper titled "Fake News: A New Benchmark Dataset" by William Yang Wang, published in 2017. The dataset is designed to facilitate the development and evaluation of algorithms for the automatic detection of fake news or misinformation.

More details about the dataset can be found in the *data.md* in the attachment.

References

Ahmed, A. A. A., Aljabouh, A., Donepudi, P. K., and Choi, M. S. (2021). Detecting fake news using machine learning : A systematic literature review.

Capuano, N., Fenza, G., Loia, V., and Nota, F. D. (2023). Content-based fake news detection with machine and deep learning: a systematic review. *Neurocomputing*, 530:91–103.

Konkobo, P. M., Zhang, R., Huang, S., Minoungou, T. T., Ouedraogo, J. A., and Li, L. (2020). A deep learning model for early detection of fake news on social media. In *2020 7th International Conference on Behavioural and Social Computing (BESC)*, pages 1–6.

Ma, J., Gao, W., Mitra, P., Kwon, S., Jansen, J., Wong, K.-F., and Cha, M. (2016). Detecting rumors from microblogs with recurrent neural networks.

Ma, J., Gao, W., and Wong, K.-F. (2019). Detect rumors on twitter by promoting information campaigns with generative adversarial learning. In *The World Wide Web Conference*, WWW '19, page 3049–3055, New York, NY, USA. Association for Computing Machinery.

Mansouri, R., Naderan-Tahan, M., and Rashti, M. J. (2020). A semi-supervised learning method for fake news detection in social media. In *2020 28th Iranian Conference on Electrical Engineering (ICEE)*, pages 1–5.

Manzoor, S. I., Singla, J., and Nikita (2019). Fake news detection using machine learning approaches: A systematic review. In *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pages 230–234.

Vaibhav, V., Mandyam, R., and Hovy, E. (2019). Do sentence interactions matter? leveraging sentence level representations for fake news classification. In Ustalov, D., Somasundaran, S., Jansen, P., Glavaš, G., Riedl, M., Surdeanu, M., and Vazirgiannis, M., editors, *Proceedings of the Thirteenth Workshop on Graph-Based Methods for Natural Language Processing (TextGraphs-13)*, pages 134–139, Hong Kong. Association for Computational Linguistics.

Yu, F., Liu, Q., Wu, S., Wang, L., and Tan, T. (2017). A convolutional approach for misinformation identification. In *International Joint Conference on Artificial Intelligence*.

[2]https://www.cs.ucsb.edu/ william/data/liar_dataset.zip