Combining Network Structures and Natural Language Processing for Classification Tasks

Gregor Donabauer

University of Regensburg, Regensburg, Germany

Abstract

Recent advances around large language models have led to significant improvements in many Information Retrieval related tasks. However, certain limitations persist in classification tasks that consider individual data points, such as text documents, in isolation. In order to overcome these limitations, linking information in network-like structures can help to model more context. This paper discusses paths around the exploration of combining graph structures and Natural Language Processing, assessing their efficacy, and examining applications in the field of IR for classification. It presents a series of experiments and their corresponding results for the use case of fake news detection. While the primary focus of this paper lies on this single use case, a brief outlook on other classification areas within IR is also presented towards the end.

Keywords

Classification, Natural Language Processing, Graph Machine Learning

1. Introduction

Natural Language Processing (NLP) has seen remarkable progress in recent years due to the invention of the Transformer architecture [1], pre-training of large language models (LLMs) [2] and recently in form of fine-tuning them via reinforcement learning from human feedback (RLHF) [3]. These advancements have had a significant impact on the field of Information Retrieval (IR) where LLMs nowadays are extensively utilized in various core IR tasks, e.g. [4, 5, 6]. Additionally, they have become highly relevant in other areas of IR that involve classification, such as fake news detection [7, 8].

These advancements in recent years have undoubtedly brought benefits to IR as a whole, and to the subfield of classification. However, limitations still exist when it comes to the formalization of classification tasks. LLMs and similar models are typically designed to handle a single type of input information, such as text. However, this approach often results in limitations when it comes to capturing the broader context surrounding text documents or input entities in a more comprehensive manner [9]. This is primarily due to a lack of knowledge regarding the relationships between data from different sources, which can naturally be represented as linked data structures (graphs) [10]. These graphs consist of nodes representing data points (which can be initially represented using models like BERT) and edges representing interactions between pairs of nodes [9]. Recently, deep learning approaches have been proposed that apply

FDIA'23: Symposium on Future Directions in Information Access, Aug 28– Sep 2, 2023, Vienna, Austria
gregor.donabauer@ur.de (G. Donabauer)

convolution to the neighborhood of nodes (based on the graph structure) to learn about entities in context of their surroundings [11].

As an example, for fake news detection this could mean that a post on social media has limited expressiveness if we only look at its text but tells us much more if we consider how it is linked to users interacting with it, e.g. [12, 13, 7]. Furthermore, we can extend this analysis by observing how propagation patterns evolve over time and how rapidly they spread [14, 15].

To address the described limitations and explore paths around the combination of network structures and NLP, we want to investigate the following research questions:

RQ1: How can we model classification problems by combining graph structures and NLP, with a specific emphasis on text data?

RQ2: What types of input signals and modeling techniques can impact the effectiveness of this approach?

RQ3: How can this paradigm be applied to classification in other domains within the field of IR?

To make sure that the content of this paper aligns with its scope, we will primarily focus on addressing RQ1 and RQ2 in the context of a single use case (fake news detection). However, we will also provide a brief outlook on RQ3 towards the end. The paper is structured as follows: we start by providing an overview of relevant research on context-based fake news detection (2), and discussing our own approaches thus far in 3. The outlook on applying these approaches to other domains will be given in 4 before finally concluding in 5.

2. Related Work

Our brief review of related work will focus on the domain of fake news detection, especially on methods that follow the context-based paradigm. While at a high level one way of distinguishing the different research directions in fake news detection is to classify them into content-based methods and context-focused approaches [16], the latter one is tapping into the direction of putting together various (social media) signals as context, unlike content-based methods that solely analyze news stories in isolation.

Many types of signals, including user engagement and dissemination networks, can be considered as context. It is worth noting that representation methods are not necessarily graph-based. For instance, in fake news detection on Twitter, user-comment relations have been utilized without modeling them as graphs [17, 18]. Context has been incorporated in diverse ways, e.g. retweet features have been combined with propagation patterns, again not as graphs [19]. Another aspect of context are images. For example, Qian et al. [20] embedded text data from news articles using BERT, and combined this information with Convolutional Neural Network (CNN) encoded images associated with those articles. Similarly, Sachan et al. [21] and Wu et al. [22] adopt a CNN to encode images and fuse this information with BERT embeddings for news classification. Other approaches involve incorporating domain-specific information in their fake news classifiers [23] or consider news media profiling and textual social media context [24]. Although these methods go beyond relying solely on news text data, they still do not aim to systematically and coherently represent the structural information between different entities.

Consequently, motivated by these limitations, graph structures and graph neural networks have gained popularity in recent years, as we will show next.

Network-based approaches are used to for example model user-article graphs and their neighborhood structure is aggregated to embed news nodes, which are subsequently classified [25]. However, many of these approaches focus on a limited number of information types, such as users and news items. In similar problem formulations, user-post interaction graphs are considered [26, 27]. Mehta et al. [28] treat fake news detection as a node classification task in a large graph comprising users, news, and publishers. However, they do not account for differences in edge types. In their subsequent work, they introduce link prediction in this graph structure to find previously unknown edges [29]. Lu and Li [30] also employ user graphs to extract social interactions using Graph Neural Networks (GNNs) and, in combination with Recurrent Neural Networks (RNNs), classify related news articles by leveraging retweet patterns. Other approaches construct large heterogeneous graphs consisting of articles, users, social media features (such as tweets), and publishers, without limiting the scope to social media context [10, 31].

A recently emerging directing in fake news detection with graph-based methods is to exploit temporal network patterns [32, 14, 15, 33]. Dou et al. [32] formulate fake news detection as a graph classification task, where user representations are derived from timeline tweet embeddings and used as node features to model news propagation paths associated with articles. Song et al. [14, 15] propose propagation-based approaches using graph-structured, temporally evolving retweet patterns to classify articles as fake or true. Even heterogeneous graph structures have recently been considered in temporal manner, however, with a limited set of node types [33].

In conclusion, graph structures provide a powerful means to incorporate diverse dimensions of information and enhance fake news detection. However, the investigation of linking multiple entity and edge types in a heterogeneous manner, is still limited. Furthermore, there has been limited exploration of text embedding methods using LLMs to represent nodes, as well as the formulation of classification based on graph versus node perspectives. Our research aims to contribute to bridging this gap.

3. Exploring Fake News Detection: Experiments and Findings

3.1. Static Graphs

To outline the initial series of experiments we conducted, we will primarily summarize the setup and findings, which have also been recently published [34]. The objective of these experiments was to assess the combination of diverse node and edge types sourced from a social media platform (compare dataset [35]) within heterogeneous graphs. Our construction method of these social media graphs represents a notably comprehensive approach with a focus on social media data in comparison to existing literature.

In detail, we construct a heterogeneous social media context graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ around each news article. Thus, each snapshot consists of a set of disjoint vertex sets $\mathcal{V} = \mathcal{V}_N \cup \mathcal{V}_T \cup \mathcal{V}_U$ where $\mathcal{V}_i \cap \mathcal{V}_j = \emptyset, \forall i \neq j$. $\mathcal{V}_N$ represents instances of news articles published by news outlets and $|\mathcal{V}_N| = 1$ for each graph. Node features for this type of vertex are obtained using BERT-base document embeddings of the textual content of the articles, i.e. $\forall v_n \in \mathcal{V}_N : v_n \in \mathbb{R}^{768}$.

All embeddings are generated using flairNLP[1]. Set $\mathcal{V}_T$ includes all tweets that refer to the news article, all tweets retweeting those posts and the latest timeline tweets of each user. In our experiments we are incrementally increasing the amount of information types in the heterogeneous graphs, thus the number of nodes in $\mathcal{V}_T$ varies depending on the experimental setup. Node features are a concatenation of textual BERT-base document embeddings, as well as retweet count and favorite count, i.e. $\forall v_t \in \mathcal{V}_T : v_t \in \mathbb{R}^{770}$. Finally, $\mathcal{V}_U$ is the set of social media user profile nodes of people related to the posts in $\mathcal{V}_T$, where features are a concatenation of BERT-base user profile description embeddings, follower count, friends count, favorites count and statuses count, i.e. $\forall v_u \in \mathcal{V}_U : v_u \in \mathbb{R}^{772}$. We also evaluate graphs with text embeddings only and leave out other features (e.g. number of favorite count) in that case. Edges connecting the nodes are satisfying constraints according to the node types they link together. More specifically, we use at most three types of edges: tweets citing news articles ($(v_t, \tau_{TN}, v_n) \in \mathcal{E} \rightarrow v_t \in \mathcal{V}_T, v_n \in \mathcal{V}_N$), users posting tweets (which also applies to users posting retweets and users posting timeline tweets) ($(v_u, \tau_{UT}, v_t) \in \mathcal{E} \rightarrow v_u \in \mathcal{V}_U, v_t \in \mathcal{V}_T$) and tweets retweeting tweets ($(v_t, \tau_{TT}, u_t) \in \mathcal{E} \rightarrow v_t \in \mathcal{V}_T, u_t \in \mathcal{V}_T$). The graphs fuse all types of information (the news article as well as a range of different social media context features) without any prior aggregation step.

We formalize fake news detection as a graph classification problem and evaluate different GNN architectures and setups with heterogeneous combinations of node types. It turns out that the combination of as much social context as possible yields the best results, on par with or even superior to state-of-the-art performance. Another important finding is that heterogeneous data structures perform significantly better compared to homogeneous graphs. Overall, these initial experiments not only validate our approach but also establish the significance of pursuing this research direction.

3.2. Temporal Graphs

As mentioned earlier, these graphs are not static in reality but instead evolve over time [14]. Therefore, instead of considering a static snapshot, we construct sequences of graphs using the methods described previously. Each snapshot within the sequence captures the interactions observable at a specific point in time after the publication of the news article. For instance, we track the evolving interactions in five-minute intervals, resulting in a continuously expanding set of context graphs. Consequently, our goal is not to classify individual graphs but rather the sequence of graphs. The baseline for each of the setups was to take the last graph of each sequence and classify them as a static snapshot which represents the same classification method as described just before. As part of these experiments we extended the state-of-the-art Python library for temporal geometric deep learning [36]. Initially, the library only supported homogeneous graph sequences. However, we integrated data structures and neural network architectures that can effectively handle heterogeneous graph sequences. In our experiments, we considered different setups. One setup focuses on one-minute intervals for a duration of 20 minutes after publication. Another setup employes five-minute intervals for a period of 120 minutes after publication. Additionally, we explore five-minute intervals for a period of 180 minutes after publication. Lastly, we use an irregular step-size as proposed by Song et al. [15].

[1] https://github.com/flairNLP/flair

It turned out that the majority of the experimental setups we evaluated did not result in improvements over the baseline. Only in one particular setup, which followed the exact temporal steps as by Song et al. [15], we observe a significant improvement. However, overall it appears that, at least for the graph structures employed in this experiment, the inclusion of temporal information does not provide as much benefit as initially anticipated.

These observations highlight the room that is left for improvements and further exploration in this domain. This refers not only to the type of constructed graphs but also to the architecture of the neural networks employed for data processing.

4. Outlook

4.1. Other Domains

An example of how these approaches can be projected to completely different domains is given at the university hospital in Regensburg. In this case, the hygiene department aims to predict outbreaks of multi-resistant germs, for example by classifying patients that have a high risk of infection. By leveraging internal databases containing patients' medical records, including information documented in medical letters, valuable insights about treatments can be extracted. Additionally, analyzing patients' movements within the hospital, specifically their transitions between wards, can provide data for constructing a network model. In essence, the same underlying characteristics mentioned earlier can be observed: dynamic network structures that evolve over time and hold information that was originally (at least partly) in textual format.

4.2. Future Work

As mentioned in the beginning, pre-training neural networks has revolutionized the field of NLP [2] and has demonstrated its effectiveness in computer vision, e.g. [37]. In a recent development, this approach has also been proposed in the context of graph machine learning [38]. Hu et al. [38] have successfully employed pre-training to learn general properties of graphs representing molecules in Chemistry and proteins in Biology. Inspired by this progress, our next objective is to apply similar techniques to learn general properties of the social media graphs described so far via pre-training. By doing so, we aim to investigate the impact on zero-shot performance of fake news detection and classification after fine-tuning these pre-trained models.

Apart from that, it could be interesting to see how reinforcement learning (which was also mentioned in the beginning [3]) could be integrated into the overall problem formulation. Wang et al. [39] recently made a first attempt to do so in context of graph-based fake news detection.

5. Conclusion

In this paper, we presented an overview of our research on classification tasks in the field of Information Retrieval, specifically focusing on the detection of fake news. While certain approaches we have explored have shown promise in improving detection performance, there remains a significant amount of work to be done. Additionally, we tried to describe our next research directions that will address the research questions initially proposed.

References

[1] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is All you Need, in: I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett (Eds.), Advances in Neural Information Processing Systems, volume 30, Curran Associates, Inc., 2017. URL: https://proceedings.neurips.cc/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

[2] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training 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, 2019, pp. 4171–4186. URL: https://aclanthology.org/N19-1423. doi:10.18653/v1/N19-1423.

[3] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray, J. Schulman, J. Hilton, F. Kelton, L. Miller, M. Simens, A. Askell, P. Welinder, P. F. Christiano, J. Leike, R. Lowe, Training language models to follow instructions with human feedback, in: S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, A. Oh (Eds.), Advances in Neural Information Processing Systems, volume 35, Curran Associates, Inc., 2022, pp. 27730–27744. URL: https://proceedings.neurips.cc/paper_files/paper/2022/file/b1efde53be364a73914f58805a001731-Paper-Conference.pdf.

[4] Q. Dong, Y. Liu, S. Cheng, S. Wang, Z. Cheng, S. Niu, D. Yin, Incorporating Explicit Knowledge in Pre-Trained Language Models for Passage Re-Ranking, in: Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '22, Association for Computing Machinery, New York, NY, USA, 2022, p. 1490–1501. URL: https://doi.org/10.1145/3477495.3531997. doi:10.1145/3477495.3531997.

[5] C. Lassance, H. Dejean, S. Clinchant, An Experimental Study on Pretraining Transformers from Scratch for IR, in: J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, A. Caputo (Eds.), Advances in Information Retrieval, Springer Nature Switzerland, Cham, 2023, pp. 504–520.

[6] M. M. Abdollah Pour, P. Farinneya, A. Toroghi, A. Korikov, A. Pesaranghader, T. Sajed, M. Bharadwaj, B. Mavrin, S. Sanner, Self-supervised Contrastive BERT Fine-tuning for Fusion-Based Reviewed-Item Retrieval, in: J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, A. Caputo (Eds.), Advances in Information Retrieval, Springer Nature Switzerland, Cham, 2023, pp. 3–17.

[7] X. Zhou, R. Zafarani, A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities, ACM Comput. Surv. 53 (2020). URL: https://doi.org/10.1145/3395046. doi:10.1145/3395046.

[8] S. Dori-Hacohen, K. Sung, J. Chou, J. Lustig-Gonzalez, Restoring Healthy Online Discourse by Detecting and Reducing Controversy, Misinformation, and Toxicity Online, in: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Machinery, New York, NY, USA, 2021, p. 2627–2628. URL: https://doi.org/10.1145/3404835.3464926.

[9] W. L. Hamilton, Graph Representation Learning, Synthesis Lectures on Artificial Intelli-

gence and Machine Learning 14 (2020) 1–159.
[10] V.-H. Nguyen, K. Sugiyama, P. Nakov, M.-Y. Kan, FANG: Leveraging Social Context for Fake News Detection Using Graph Representation, in: Proceedings of the 29th ACM International Conference on Information & Knowledge Management, CIKM '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 1165–1174. URL: https://doi.org/10.1145/3340531.3412046. doi:10.1145/3340531.3412046.
[11] T. N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, in: International Conference on Learning Representations (ICLR 2017), 2017.
[12] K. Shu, S. Wang, H. Liu, Beyond News Contents: The Role of Social Context for Fake News Detection, in: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, WSDM '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 312–320. URL: https://doi.org/10.1145/3289600.3290994. doi:10.1145/3289600.3290994.
[13] E. Mosca, M. Wich, G. Groh, Understanding and Interpreting the Impact of User Context in Hate Speech Detection, in: Proceedings of the Ninth International Workshop on Natural Language Processing for Social Media, Association for Computational Linguistics, Online, 2021, pp. 91–102. URL: https://aclanthology.org/2021.socialnlp-1.8. doi:10.18653/v1/2021.socialnlp-1.8.
[14] C. Song, K. Shu, B. Wu, Temporally evolving graph neural network for fake news detection, Information Processing & Management 58 (2021) 102712. URL: https://www.sciencedirect.com/science/article/pii/S0306457321001965. doi:https://doi.org/10.1016/j.ipm.2021.102712.
[15] C. Song, Y. Teng, Y. Zhu, S. Wei, B. Wu, Dynamic graph neural network for fake news detection, Neurocomputing 505 (2022) 362–374. URL: https://www.sciencedirect.com/science/article/pii/S0925231222009158. doi:https://doi.org/10.1016/j.neucom.2022.07.057.
[16] K. Sharma, F. Qian, H. Jiang, N. Ruchansky, M. Zhang, Y. Liu, Combating Fake News: A Survey on Identification and Mitigation Techniques, ACM Trans. Intell. Syst. Technol. 10 (2019). URL: https://doi.org/10.1145/3305260. doi:10.1145/3305260.
[17] K. Shu, L. Cui, S. Wang, D. Lee, H. Liu, DEFEND: Explainable Fake News Detection, in: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, KDD '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 395–405. URL: https://doi.org/10.1145/3292500.3330935. doi:10.1145/3292500.3330935.
[18] S. Raza, Automatic Fake News Detection in Political Platforms - A Transformer-based Approach, in: Proceedings of the 4th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE 2021), Association for Computational Linguistics, Online, 2021, pp. 68–78. URL: https://aclanthology.org/2021.case-1.10. doi:10.18653/v1/2021.case-1.10.
[19] Y. Liu, Y.-F. Wu, Early Detection of Fake News on Social Media Through Propagation Path Classification with Recurrent and Convolutional Networks, Proceedings of the AAAI Conference on Artificial Intelligence 32 (2018). URL: https://ojs.aaai.org/index.php/AAAI/article/view/11268.
[20] S. Qian, J. Wang, J. Hu, Q. Fang, C. Xu, Hierarchical Multi-Modal Contextual At-

tention Network for Fake News Detection, in: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Machinery, New York, NY, USA, 2021, p. 153–162. URL: https://doi.org/10.1145/3404835.3462871.

[21] T. Sachan, N. Pinnaparaju, M. Gupta, V. Varma, SCATE: Shared Cross Attention Transformer Encoders for Multimodal Fake News Detection, in: Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 399–406. URL: https://doi.org/10.1145/3487351.3490965. doi:10.1145/3487351.3490965.

[22] Y. Wu, P. Zhan, Y. Zhang, L. Wang, Z. Xu, Multimodal Fusion with Co-Attention Networks for Fake News Detection, in: Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, Association for Computational Linguistics, Online, 2021, pp. 2560–2569.

[23] Q. Nan, J. Cao, Y. Zhu, Y. Wang, J. Li, MDFEND: Multi-Domain Fake News Detection, in: Proceedings of the 30th ACM International Conference on Information & Knowledge Management, Association for Computing Machinery, New York, NY, USA, 2021, p. 3343–3347. URL: https://doi.org/10.1145/3459637.3482139.

[24] R. Baly, G. Karadzhov, J. An, H. Kwak, Y. Dinkov, A. Ali, J. Glass, P. Nakov, What Was Written vs. Who Read It: News Media Profiling Using Text Analysis and Social Media Context, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Online, 2020, pp. 3364–3374. URL: https://aclanthology.org/2020.acl-main.308. doi:10.18653/v1/2020.acl-main.308.

[25] S. Chandra, P. Mishra, H. Yannakoudakis, M. Nimishakavi, M. Saeidi, E. Shutova, Graph-based Modeling of Online Communities for Fake News Detection, CoRR abs/2008.06274 (2020). URL: https://arxiv.org/abs/2008.06274. arXiv:2008.06274.

[26] E. Min, Y. Rong, Y. Bian, T. Xu, P. Zhao, J. Huang, S. Ananiadou, Divide-and-Conquer: Post-User Interaction Network for Fake News Detection on Social Media, in: Proceedings of the ACM Web Conference 2022, WWW '22, Association for Computing Machinery, New York, NY, USA, 2022, p. 1148–1158. URL: https://doi.org/10.1145/3485447.3512163. doi:10.1145/3485447.3512163.

[27] S. Rode-Hasinger, A. Kruspe, X. X. Zhu, True or False? Detecting False Information on Social Media Using Graph Neural Networks, in: Proceedings of the Eighth Workshop on Noisy User-generated Text (W-NUT 2022), Association for Computational Linguistics, Gyeongju, Republic of Korea, 2022, pp. 222–229. URL: https://aclanthology.org/2022.wnut-1.24.

[28] N. Mehta, D. Goldwasser, Tackling Fake News Detection by Interactively Learning Representations using Graph Neural Networks, in: Proceedings of the First Workshop on Interactive Learning for Natural Language Processing, Association for Computational Linguistics, Online, 2021, pp. 46–53. URL: https://aclanthology.org/2021.internlp-1.7. doi:10.18653/v1/2021.internlp-1.7.

[29] N. Mehta, M. Pacheco, D. Goldwasser, Tackling Fake News Detection by Continually Improving Social Context Representations using Graph Neural Networks, in: Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Dublin, Ireland, 2022, pp. 1363–1380. URL: https://aclanthology.org/2022.acl-long.97.

[30] Y.-J. Lu, C.-T. Li, GCAN: Graph-aware Co-Attention Networks for Explainable Fake News Detection on Social Media, in: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Online, 2020, pp. 505–514. URL: https://aclanthology.org/2020.acl-main.48. doi:10.18653/v1/2020.acl-main.48.

[31] T. Li, Y. Sun, S.-l. Hsu, Y. Li, R. C.-W. Wong, Fake News Detection with Heterogeneous Transformer, 2022. URL: https://arxiv.org/abs/2205.03100. doi:10.48550/ARXIV.2205.03100.

[32] Y. Dou, K. Shu, C. Xia, P. S. Yu, L. Sun, User Preference-Aware Fake News Detection, in: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, Association for Computing Machinery, New York, NY, USA, 2021, p. 2051–2055. URL: https://doi.org/10.1145/3404835.3462990.

[33] J. H. Go, A. Sari, J. Jiang, S. Yang, S. Jha, Fake News Quick Detection on Dynamic Heterogeneous Information Networks, 2022. arXiv:2205.07039.

[34] G. Donabauer, U. Kruschwitz, Exploring Fake News Detection with Heterogeneous Social Media Context Graphs, in: J. Kamps, L. Goeuriot, F. Crestani, M. Maistro, H. Joho, B. Davis, C. Gurrin, U. Kruschwitz, A. Caputo (Eds.), Advances in Information Retrieval, Springer Nature Switzerland, Cham, 2023, pp. 396–405.

[35] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, H. Liu, FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media, Big Data 8 (2020) 171–188. URL: https://doi.org/10.1089/big.2020.0062. doi:10.1089/big.2020.0062, pMID: 32491943.

[36] B. Rozemberczki, P. Scherer, Y. He, G. Panagopoulos, A. Riedel, M. Astefanoaei, O. Kiss, F. Beres, G. López, N. Collignon, R. Sarkar, PyTorch Geometric Temporal: Spatiotemporal Signal Processing with Neural Machine Learning Models, in: Proceedings of the 30th ACM International Conference on Information & Knowledge Management, CIKM '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 4564–4573. URL: https://doi.org/10.1145/3459637.3482014. doi:10.1145/3459637.3482014.

[37] K. He, R. Girshick, P. Dollar, Rethinking ImageNet Pre-Training, in: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019.

[38] W. Hu, B. Liu, J. Gomes, M. Zitnik, P. Liang, V. Pande, J. Leskovec, Strategies for Pre-training Graph Neural Networks, in: International Conference on Learning Representations (ICLR 2020), 2020.

[39] H. Wang, Y. Dou, C. Chen, L. Sun, P. S. Yu, K. Shu, Attacking Fake News Detectors via Manipulating News Social Engagement, in: Proceedings of the ACM Web Conference 2023, WWW '23, Association for Computing Machinery, New York, NY, USA, 2023, p. 3978–3986. URL: https://doi.org/10.1145/3543507.3583868. doi:10.1145/3543507.3583868.