$5^{th}$ International Conference on Computational Social Science $IC^2S^2$
July 17-20, 2019, University of Amsterdam, The Netherlands

# Semi-Supervised Learning and Graph Neural Networks for Fake News Detection

Adrien Benamira[1], Benjamin Devillers[1], Etienne Lesot[1], Ayush K. Rai[1], Manal Saadi[1], and Fragkiskos D. Malliaros[1,2]

[1]CentraleSupélec, University of Paris-Saclay, France [2]Inria Saclay, France

## Extended Abstract

Social networks have become the main platforms for information dissemination. Nevertheless, due to the increasing number of users, social media platforms tend to be highly vulnerable to the propagation of disinformation – making the detection of fake news a challenging task. In this work, we focus on content-based methods for detecting fake news – casting the problem to a binary text classification one (an article corresponds to either fake news or not). The main challenge here stems from the fact that the number of labeled data is limited; very few articles can be checked and annotated as fake. To this extend, we opted for semi-supervised learning approaches [1]. In particular, our work proposes a graph-based semi-supervised fake news detection method, based on graph neural networks.

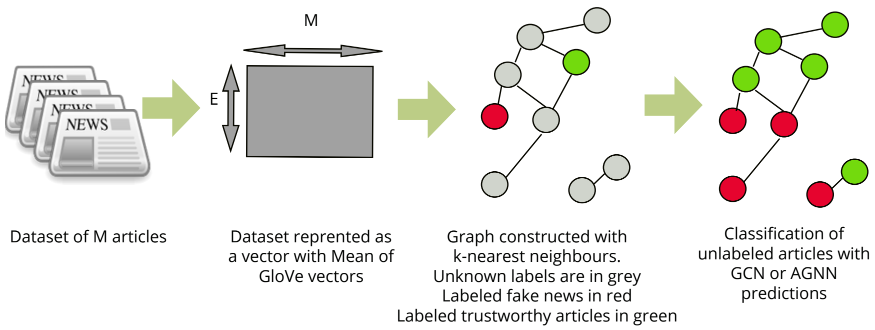


Figure 1: Illustration of the proposed approach: $M$ denotes the number of articles (real and fake) and $E$ is the dimension of our GloVe embeddings (in our case, $M = 150$, $E = 100$). Finally, we use $k = 4$ nearest neighbors to build the graph.

A schematic representation of our approach is depicted in Figure 1. Considering a collection of $M$ articles, our graph-based method is composed by the following three components.

**Embedding of articles:** Here we are considering a vector representation of an article. Our approach is based on pre-trained GloVe word embeddings [5], computing the mean vector of the words appearing within an article.

**Graph construction:** The second step concerns the construction of similarity graph among articles [1]. In particular, for each article (i.e., node in the graph), we look for the $k$–nearest neighbours based on the Euclidean distances in the embedding space.

**Classification:** For the classification task over the similarity graph between articles, we use two neural network graph learning methods: Graph Convolutional Networks (GCN) [3] and Attention Graph Neural Network (AGNN). [6]

**Empirical Evaluation.** In our preliminary empirical evaluation, we have compared our graph neural network-based methods against the approach by Guacho et al. [1], which follows a similar framework. In that method, the embedding of articles is obtained using CP/PARAFAC tensor decomposition on the binary co-occurrence matrix between all articles, and the classification is performed using the Fast Belief Propagation (FaBP) algorithm [4]. We evaluate our method on a recent dataset presented in [2], which is comprised of 150 labeled articles, 75 of those are fake news and 75 real. We pick the labeled articles at random and average the results over 20 tries. For the neural graph networks, we use 4 layers, 4 neighbours, 16 hidden units, a learning rate of 0.01 and a weight decay of $5e-4$. We train our graph neural networks during 1000 epochs and we keep the one which has the best accuracy on the test (unlabeled) data.

| Methods | **Accuracy (in %)** | | | | |
|---|---|---|---|---|---|
| | *2 % labeled data* | *5 % labeled data* | *10 % labeled data* | *15 % labeled data* | 20 % labeled data |
| Guacho et al. [1] | $56.65 \pm 9.67$ | $63.60 \pm 7.52$ | $70.95 \pm 5.28$ | $74.05 \pm 3.80$ | $79.8 \pm 3.10$ |
| Ours with AGNN | $70.45 \pm 5.39$ | $72.00 \pm 8.05$ | $78.70 \pm 3.54$ | $\mathbf{83.35 \pm 1.74}$ | $84.25 \pm 3.51$ |
| Ours with GCN | $\mathbf{72.04 \pm 6.00}$ | $\mathbf{77.35 \pm 3.72}$ | $\mathbf{79.85 \pm 3.41}$ | $82.35 \pm 2.44$ | $\mathbf{84.94 \pm 2.30}$ |

Table 1: Classification accuracy under different fractions of labeled data used for training.

Table 1 gives the classification accuracy of the different methods, varying the amount of labeled data used for training (ranging from 2% up to 20%). As we can observe, the proposed graph neural network approaches achieve a performance improvement of up to 19% with only 10% of the labeled data, while they are more stable reducing the standard deviation of the results. Furthermore, our AGNN and GCN methods are computationally faster when it comes to evaluate a new article. We are currently working to further extend our study by considering more baseline methods and testing the performance on bigger as well as multi-labeled fake news datasets.

# References

[1] Gisel Bastidas Guacho, Sara Abdali, Neil Shah, and Evangelos E. Papalexakis. Semi-supervised content-based detection of misinformation via tensor embeddings. In *ASONAM*, pages 322–325, 2018.

[2] Benjamin D. Horne and Sibel 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. *arXiv*, 2017.

[3] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *ICLR*, 2017.

[4] Danai Koutra, Tai-You Ke, U Kang, Duen Horng Polo Chau, Hsing-Kuo Kenneth Pao, and Christos Faloutsos. Unifying guilt-by-association approaches: Theorems and fast algorithms. In *ECML-PKDD*, pages 245–260, 2011.

[5] Matthew E Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. Deep contextualized word representations. In *NAACL-HLT*, pages 2227–2237, 2018.

[6] Kiran K. Thekumparampil, Chong Wang, Sewoong Oh, and Li-Jia Li. Attention-based graph neural network for semi-supervised learning. In *ICLR*, 2018.