MeVer Team Tackling Corona Virus and 5G Conspiracy Using Ensemble Classification Based on BERT

Olga Papadopoulou
Information Technologies Institute, CERTH, Thessaloniki, Greece
olgapapa@iti.gr

Giorgos Kordopatis-Zilos
Information Technologies Institute, CERTH, Thessaloniki, Greece
georgekordopatis@iti.gr

Symeon Papadopoulos
Information Technologies Institute, CERTH, Thessaloniki, Greece
papadop@iti.gr

ABSTRACT

This paper presents the approach developed by the Media Verification (MeVer) team to tackle the task of FakeNews: Coronavirus and 5G conspiracy at the MediaEval 2020 Challenge. We build a two-stage classification approach based on ensemble learning of multi-

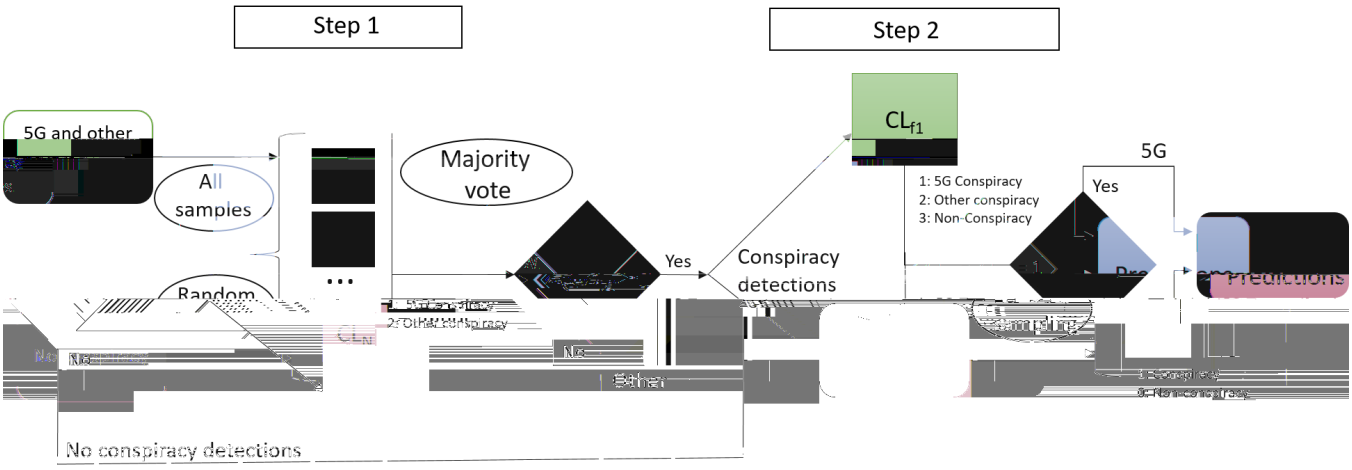

Figure 1: Our proposed pipeline for tackling the challenge of Corona virus and 5G conspiracy

Table 1: Summary of the training samples used to build the respective models

Label	CL_{8}	$CL_{<D;C8}$	$CL_{2>=B?}$
5G conspiracy	1847	712	712
Other conspiracy		712	712
Non-conspiracy	1847	712	-

In the second step, the predictions of *Non-conspiracy* are considered as final predictions without further processing while the *Conspiracy* tweets are further processed to distinguish *5G-conspiracy* from *Other-conspiracy*. In this step, two additional models are trained focusing on the detection of *5G-conspiracy* tweets. The first, $CL_{5\,1}$, is a three-class model (1: *5G-conspiracy*, 2: *Other-conspiracy* and 3: *Non-conspiracy*) trained using random samples from the majority classes and the total number of minority class samples (*Other-conspiracy*). The other model, $CL_{5\,2}$, is a binary classifier trained on the two *Conspiracy* classes. The final decision is taken if $CL_{5\,1}$ = $CL_{5\,2}$ = 1 = *5G-conspiracy*. In any other case, the tweet is labeled as *Other-conspiracy*.

3.1 Implementation details

For tokenization, we employ `bert-base-uncased` of BertTokenizer applied to the text of the tweets. The text is limited to 160 tokens as input to the network. Considering that the maximum tweet length is 280 characters, it is most likely that the entire text is processed to calculate the prediction. As a backbone network, we employ the `bert-base-uncased` version of BERT [13], which is a compact transformer model, trained on lower-cased English text. The network architecture consists of 12 layers (i.e., Transformer blocks), with 768 hidden units, and 12 heads for multi-head attention layers, resulting in a total of 109M parameters.

REFERENCES

[1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805* (2018).

[2] Mohamed K Elhadad, Kin Fun Li, and Fayez Gebali. 2020. Detecting Misleading Information on COVID-19. *IEEE Access* 8 (2020), 165201–165215.

[3] Tamanna Hossain, Robert L Logan IV, Arjuna Ugarte, Yoshitomo Matsubara, Sameer Singh, and Sean Young. 2020. Detecting covid-19 misinformation on social media. (2020).

[4] Justin M Johnson and Taghi M Khoshgoftaar. 2019. Survey on deep learning with class imbalance. *Journal of Big Data* 6, 1 (2019), 27.

[5] Heejung Jwa, Dongsuk Oh, Kinam Park, Jang Mook Kang, and Heuiseok Lim. 2019. exBAKE: Automatic Fake News Detection Model Based on Bidirectional Encoder Representations from Transformers (BERT). *Applied Sciences* 9, 19 (2019), 4062.

[6] Diederik P Kingma and Jimmy Ba. 2014. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

[7] Chao Liu, Xinghua Wu, Min Yu, Gang Li, Jianguo Jiang, Weiqing Huang, and Xiang Lu. 2019. A Two-Stage Model Based on BERT for Short Fake News Detection. In *International Conference on Knowledge Science, Engineering and Management*. Springer, 172–183.

[8] Olga Papadopoulou, Markos Zampoglou, Symeon Papadopoulos, and Ioannis Kompatsiaris. 2019. A corpus of debunked and veri□ed user-generated videos. *Online information review* (2019).

[9] Konstantin Pogorelov, Daniel Thilo Schroeder, Luk Burchard, Johannes Moe, Stefan Brenner, Petra Filkukova, and Johannes Langguth. 2020. FakeNews: Corona Virus and 5G Conspiracy Task at MediaEval 2020. In *MediaEval 2020 Workshop*.

[10] Juan Carlos Medina Serrano, Orestis Papakyriakopoulos, and Simon Hegelich. 2020. NLP-based Feature Extraction for the Detection of COVID-19 Misinformation Videos on YouTube. In *Proceedings of the 1st Workshop on NLP for COVID-19 at ACL 2020*.

[11] Karishma Sharma, Sungyong Seo, Chuizheng Meng, Sirisha Rambhatla, Aastha Dua, and Yan Liu. 2020. Coronavirus on social media: Analyzing misinformation in Twitter conversations. *arXiv preprint arXiv:2003.12309* (2020).

[12] Samia Tasnim, Md Mahbub Hossain, and Hoimonty Mazumder. 2020. Impact of rumors or misinformation on coronavirus disease (COVID-19) in social media. (2020).

[13] Iulia Turc, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Well-read students learn better: On the importance of pre-training compact models. *arXiv preprint arXiv:1908.08962* (2019).

[14] Tom Warren. 2020. British 5G towers are being set on □re because of coronavirus conspiracy theories. (Apr 2020). https://www.theverge.com/2020/4/4/21207927/5g-towers-burning-uk-coronavirus-conspiracy-theory-link

[15] Thomas Wolf, Lysandre Debut, Victor Sanh, Julien Chaumond, Clement Delangue, Anthony Moi, Pierric Cistac, Tim Rault, Rémi Louf, Morgan Funtowicz, Joe Davison, Sam Shleifer, Patrick von Platen, Clara Ma, Yacine Jernite, Julien Plu, Canwen Xu, Teven Le Scao, Sylvain Gugger, Mariama Drame, Quentin Lhoest, and Alexander M. Rush. 2020. Transformers: State-of-the-Art Natural Language Processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Association for Computational Linguistics, Online, 38–45. https://www.aclweb.org/anthology/2020.emnlp-demos.6