



g2tmn at Constraint@AAAI2021: Exploiting CT-BERT and Ensembling Learning for COVID-19 Fake News Detection

Anna Glazkova¹(✉)() , Maksim Glazkov²() , and Timofey Trifonov^{1,2}()

¹ University of Tyumen, ul. Volodarskogo 6, 625003 Tyumen, Russia
a.v.glazkova@utmn.ru

² “Organization of Cognitive Associative Systems” LLC, ul. Gertsena 64,
625000 Tyumen, Russia

Abstract. The COVID-19 pandemic has had a huge impact on various areas of human life. Hence, the coronavirus pandemic and its consequences are being actively discussed on social media. However, not all social media posts are truthful. Many of them spread fake news that cause panic among readers, misinform people and thus exacerbate the effect of the pandemic. In this paper, we present our results at the Constraint@AAAI2021 Shared Task: COVID-19 Fake News Detection in English. In particular, we propose our approach using the transformer-based ensemble of COVID-Twitter-BERT (CT-BERT) models. We describe the models used, the ways of text preprocessing and adding extra data. As a result, our best model achieved the weighted F1-score of 98.69 on the test set (the first place in the leaderboard) of this shared task that attracted 166 submitted teams in total.

Keywords: COVID-Twitter-BERT · Social media · Fake news · Ensembling learning · Coronavirus · Infodemic · Text classification

1 Introduction

Social media is a unique source of information. On the one hand, their low cost, easy access and distribution speed make it possible to quickly share the news. On the other hand, the quality and reliability of social media news is difficult to verify [38]. This is the source of a lot of false information that has a negative impact on society.

Over the past year, the world has been watching the situation developing around the novel coronavirus pandemic. The COVID-19 pandemic has become a significant newsworthy event of 2020. Therefore, news related to COVID-19 are actively discussed on social media and this topic generates a lot of misinformation. Fake news related to the pandemic have large-scale negative social consequences, they provoke huge public rumor spreading and misunderstanding about

the COVID-19 and aggravate effects of the pandemic. Moreover, recent studies [22] show an increase in symptoms such as anxiety and depression in connection with the pandemic. This is closely related to the spread of misinformation, because fake news can be more successful when the population is experiencing a stressful psychological situation [25]. The popularity of fake news on social media can rapidly increase, because the rebuttal is always published too late. In this regard, there is evidence that the development of tools for automatic COVID-19 fake news detection plays a crucial role in the regulation of information flows.

In this paper, we present our approach for the Constraint@AAAI2021 Shared Task: COVID-19 Fake News Detection in English [29] that attracted 433 participants on CodaLab. This approach achieved the weighted F1-score of 98.69 (the first place in the leaderboard) on the test set among 166 submitted teams in total.

The rest of the paper is organized as follows. A brief review of related work is given in Sect. 2. The definition of the task has been summarized in Sect. 3, followed by a brief description of the data used in Sect. 4. The proposed methods and experimental settings have been elaborated in Sect. 5. Section 6 contains the results and error analysis respectively. Section 7 is a conclusion.

2 Related Work

In recent years, the task of detecting fake news and rumors is extremely relevant. False information spreading involves various research tasks, including: fact checking [4, 40], topic credibility [15, 41], fake news spreaders profiling [34], and manipulation techniques detection [8]. Various technologies and approaches in this field range from traditional machine learning methods [5, 23, 33], to state-of-the-art transformers [24, 47].

A overview of fake news detection approaches and challenges on social media has been discussed in [38, 50]. Many scholars have proposed their solutions to this problem in different subject areas (in particular, [6, 35]). Up to now, a large number of studies in fake news detection have used supervised methods including models based on transformers-based architecture [13, 17, 49].

Some recent work have focused on detecting fake news about COVID-19. For example, the predictors of the sharing of false information about the pandemic are discussed in [3]. In [44], a novel COVID-19 fact checking algorithm is proposed that retrieves the most relevant facts concerning user claims about particular facts. A number of studies have begun to examine COVID-19 fake news detection methods for non-English languages [10, 14, 48].

In addition, several competitions have been announced over the past year related to the analysis of posts about COVID-19 on social media [1, 27, 36].

3 Task Definition

The task focused on the detection of COVID-19-related fake news in English. The sources of data was various social-media platforms such as Twitter, Facebook, Instagram, etc. Formally, the task is described as follows.

- **Input.** Given a social media post.
- **Output.** One of two different labels, such as “fake” or “real”.

The official competition metric was F1-score averaged across the classes (the weighted F1-score). The participants were allowed five submissions per team throughout the test phase.

4 Dataset

The dataset [28] provided to the participants of the shared task contains 10,700 manually annotated social media posts divided into training (6420), validation (2140), and test (2140) sets. The vocabulary size (i.e., unique words) of the dataset is 37,505 with 5141 common words in both fake and real news. The dataset contains the post ID, the post, and the corresponding label which is “fake” or “real” (see Table 1).

Table 1. Some examples of fake and real posts

Label	Post
real	The CDC currently reports 99031 deaths. In general the discrepancies in death counts between different sources are small and explicable. The death toll stands at roughly 100000 people today
real	#IndiaFightsCorona: We have 1524 #COVID testing laboratories in India and as on 25th August 2020 36827520 tests have been done: @ProfBhargava DG @ICMRDELHI #StaySafe #IndiaWillWin https://t.co/Yh3ZxknnhZ
fake	Politically Correct Woman (Almost) Uses Pandemic as Excuse Not to Reuse Plastic Bag https://t.co/thF8GuNFPe #coronavirus #nashville
fake	Obama Calls Trump’s Coronavirus Response A Chaotic Disaster https://t.co/DeDqZEhAsB

5 Our Approach

In this section, we describe the approaches that we evaluated on the validation data during the validation phase. We used transformer-based models as they demonstrate state-of-the-art results in most text classification tasks. We also evaluated the empirical effectiveness of a Linear Support Vector baseline (Linear SVC) and different text preprocessing techniques and adding extra data. The results are shown in the next section.

5.1 Data Preprocessing

Our approaches to text preprocessing for transformer-based models are various combinations of the following steps, most of which have been inspired by [18, 42]:

- removing or tokenizing hashtags, URLs, emoji, and mentions using a pre-processing library for tweet data written in Python [43]. Tokenization means the replacement of URLs, mentions, and emoji with special tokens, such as \$URL\$, \$MENTION\$, and \$HASHTAG\$ respectively (for example “HHS to distribute \$4 billion to #COVID hot spots; \$340 million already paid out. <https://t.co/uAj29XA1Y5>” (original) → “HHS to distribute \$4 billion to \$HASHTAG\$ hot spots; \$340 million already paid out. \$URL\$” (tokenizing); “HHS to distribute \$4 billion to hot spots; \$340 million already paid out.” (removing));
- using the Python emoji library to replace the emoji with short textual description [11], for example :red_heart:, :thumbs_up:, etc.;
- converting hashtags to words (“#COVID” → “COVID”);
- translating in the lowercase.

In the case of the baseline, we translated the text to the lowercase, removed punctuation and special characters, and then lemmatized the words. Further, we converted texts into the form of a token counts matrix (a bag of words model).

5.2 Models

We experimented with the following transformer-based models:

- **BERT** [9]. BERT is a language representation model presented by Google, which stands for Bidirectional Encoder Representations from Transformers. BERT-based models show great results in many natural language processing tasks. In our work, we used BERT-base-uncased, which is pretrained on texts from Wikipedia.
- **RoBERTa** [19]. RoBERTa is a robustly optimized BERT approach introduced at Facebook. Unlike BERT, RoBERTa removes the Next Sentence Prediction task from the pretraining process. RoBERTa also uses larger batch sizes and dynamic masking so that the masked token changes while training instead of the static masking pattern used in BERT. We experimented with RoBERTa-large.
- **COVID-Twitter-BERT** [26]. CT-BERT is a transformer-based model, pretrained on a large corpus of Twitter messages on the topic of COVID-19 collected during the period from January 12 to April 16, 2020. CT-BERT is optimised to be used on COVID-19 content, in particular social media posts from Twitter. This model showed a 10–30% marginal improvement compared to its base model, BERT-large, on five different specialised datasets. Moreover, it was successfully used for a variety of natural language tasks, such as identification of informative COVID-19 tweets [18], sentiment analysis [16], and claims verification [2, 45].

5.3 Additional Data

To improve the quality of our approach, we made attempts to add extra data to the model. For this purpose we used two datasets related to the topic of COVID-19 fake news:

- **CoAID: COVID-19 Healthcare Misinformation Dataset** [7]. The dataset includes 4251 health-related fake news posted on websites and social platforms.
- **FakeCovid - A Multilingual Cross-domain Fact Check News Dataset for COVID-19** [37]. The dataset contains 5182 fact-checked news articles for COVID-19 collected from January to May 2020.

In our experiments, we added news headlines to the training set.

5.4 Experimental Settings

We conducted our experiments on Google Colab Pro (CPU: Intel(R) Xeon(R) CPU @ 2.20GHz; RAM: 25.51 GB; GPU: Tesla P100-PCIE-16 GB with CUDA 10.1). Each model was trained on the training set for 3 epochs and evaluated on the validation set. As our resources are constrained, we used random seeds to fine-tune pre-trained language models and made attempts to combine them with other parameters. The models are optimised using AdamW [21] with a learning rate of $2e-5$ and epsilon of $1e-8$, max sequence length of 128 tokens, and a batch size of 8. We implemented our models using Pytorch [30] and Huggingface’s Transformers [46] libraries.

The Linear SVC was implemented with Scikit-learn [31]. For text preprocessing, we used NLTK [20] and Scikit-learn’s CountVectorizer with a built-in list of English stop-words and a maximum feature count of 10,000.

6 Results and Discussion

6.1 Comparison of Models for Fake News Detection

In Table 2, we present the results of our experiments in a step by step manner. We started with a Linear SVC baseline and then evaluated BERT-based models using a variety of text preprocessing and adding extra data techniques. Note that we evaluated our models using F1-score for the fake class while the official competition metric was the weighted F1-score.

Table 2. Evaluation results

Model	Data preprocessing	Additional data	F1-score (% for fake class)
LinearSVC	Converting into a bag of words	No	88.39
BERT	Lowercase	No	96.75
RoBERTa	Lowercase	No	97.62
RoBERTa	Removing hashtags, URLs, emoji + lowercase	No	95.79
RoBERTa	Removing URLs and emoji + converting hashtags to words + lowercase	No	95.68
CT-BERT	Lowercase	No	98.07
CT-BERT	Tokenizing URLs and mentions + converting emoji to words + lowercase	No	97.87
CT-BERT	Converting emoji to words + lowercase	No	98.32
CT-BERT	Tokenizing URLs + converting emoji to words + lowercase	No	98.42
CT-BERT	Tokenizing URLs + converting emoji to words + lowercase	FakeCovid	98.23
CT-BERT	Tokenizing URLs + converting emoji to words + lowercase	CoAID	98.37

As can be seen from the table above, CT-BERT models showed absolutely better results compared to BERT- and RoBERTa-based classifiers. Our work doesn't contain a detailed comparative analysis of text preprocessing techniques for this task. Still we can see that text preprocessing can affect the quality of fake news detection. For example, there was no evidence that removing emoji and mentions improve the model results. A clear benefit of converting hashtags into words could not be identified during this evaluation. Also, as a result of our experiments, we decided not to tokenize links to other user's accounts (mentions). This can be seen in the case of mentions of major news channels like CNN that tend to indicate that the post is real. The next section of the model evaluation was concerned with using additional datasets. We noticed that adding extra data did not show any benefits in our experiments.

6.2 Final Submissions

As it was mentioned above, the participants of the shared task were allowed five submissions per team throughout the test phase. Our best model based on experimental results (Subsect. 5.1) included the following preprocessing steps: tokenizing URLs, converting emoji to words, and lowercase. As final submissions, we used the results of hard voting ensembles of three such models with random seed values and with different data splitting into training and validation samples. The final architecture of our solution is shown in Fig. 1.

Four of our five submitted models were trained entirely on the dataset provided by the competition organizers [28]. The last model was trained on the competition data and on additional datasets [7,37]. The training details and the results of our final submissions are summarised in Table 3.

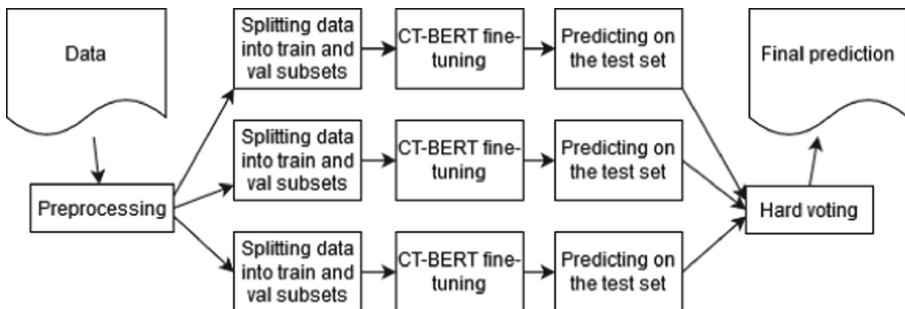
Table 3. Final submissions

Place among all submissions	Submission name	Weighted F1-score (%)	Training details
1	g2tmn_2.csv	98.69	All models were trained both on training and validation sets with no hold out validation. We trained 3 models and then used hard-voting to ensemble their predictions together
6	g2tmn_4.csv	98.51	1000 random posts were used for hold-out validation. Models were trained on all other data. We trained 5 models with random seed values and choose 3 models with the best F1-score performances to ensemble them together
11	g2tmn_1.csv	98.37	Models were trained on the official training set. The validation set was used for hold-out validation. We trained 5 models with random seed values and choose 3 best-performance models to ensemble their predictions
15	g2tmn_3.csv	98.32	This submission was similar to g2tmn_1.csv but it had different seed values
25	g2tmn_5.csv	98.18	1000 random posts were used for hold-out validation. Models were trained on all other official data, CoAID and FakeCovid data. We trained 5 models with random seed values and used 3 best-performance models for ensembling learning

It can be seen from the data in Table 3 that, with the weighted F1-score, our model performance is 98.69% (the random seeds are 23, 30, and 42), which was ranked the first place of the leaderboard of this task.

6.3 Error Analysis

Error analysis allows us to further evaluate the quality of the machine learning model and conduct a quantitative analysis of errors. Figure 2 provides the confusion matrix for our best solution when detecting fake news about COVID-19 on the test set. As can be seen from the figure, the precision of our system is slightly higher than its recall. In other words, the false negative value is greater than false positive. Table 4 shows the examples of false negative and false positive errors.

**Fig. 1.** Our approach to COVID-19 fake news detection.

We noticed that the type of error is frequently related to the topic of the post. For example, the model often misclassifies false reports about the number of people infected. At the same time, true posts related to the coronavirus vaccine or to political topics can be identified as false.

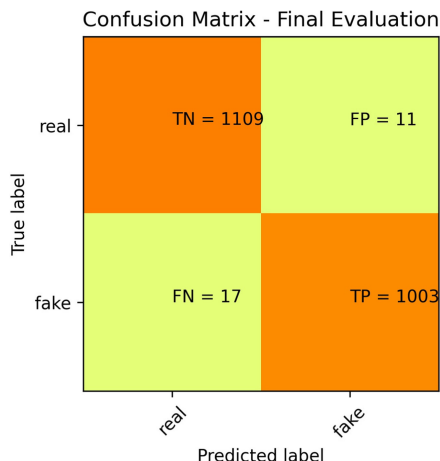


Fig. 2. Confusion matrix of our best-performance ensemble for COVID-19 fake news detection (for the fake class).

Table 4. Some examples of misclassified posts

True label	Prediction	Post
real	fake	Scientists ask: Without trial data how can we trust Russia's #COVID vaccine? https://t.co/gArcUf0Pji https://t.co/0bdcA7lf56
real	fake	*DNA Vaccine: injecting genetic material into the host so that host cells create proteins that are similar to those in the virus against which the host then creates antibodies
real	fake	Donald Trump has claimed he “up-played” the seriousness of the coronavirus pandemic - despite admitting earlier this year he had “wanted to always play it down” https://t.co/wEgmnZzrNW
fake	real	Govt has added #Corona disease in all existing mediclaim insurances as a special case #COVID2019India https://t.co/39vpW7tBqq
fake	real	As tuberculosis shaped modernism, so COVID-19 and our collective experience of staying inside for months on end will influence architecture's near future, @chaykak writes. https://t.co/ag34yZckbU
fake	real	Northern Ireland was testing for COVID-19 at a rate 10 times that of Scotland reported on 9 May 2020

7 Conclusion

In this work, we propose a simple but effective approach to COVID-19 fake news detection based on CT-BERT and ensembling learning. Our experiments confirmed that BERT-based models specialized in the subject area successfully cope with such tasks and perform high-quality binary classification.

The experimental results showed that our solution achieved 98.69% of the weighted F1-score on test data and ranked in the first place in the Constraint@AAAI2021 shared task. For future work, we can experiment with different training and data augmentation techniques. We can also apply and evaluate hybrid models combining BERT-based architectures with other methods of natural language processing [32, 39].

References

1. Alam, F., et al.: Fighting the COVID-19 infodemic: modeling the perspective of journalists, fact-checkers, social media platforms, policy makers, and the society. arXiv preprint [arXiv:2005.00033](https://arxiv.org/abs/2005.00033) (2020)
2. Alkhalifa, R. et al.: QMUL-SDS at CheckThat! 2020: determining COVID-19 tweet check-worthiness using an enhanced CT-BERT with numeric expressions. arXiv preprint [arXiv:2008.13160](https://arxiv.org/abs/2008.13160) (2020)
3. Apuke, O.D., Omar, B.: Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics Inform.* **56**, 101475 (2020)
4. Elsayed, T., et al.: Overview of the CLEF-2019 CheckThat! lab: automatic identification and verification of claims. In: Crestani, F., et al. (eds.) *CLEF 2019*. LNCS, vol. 11696, pp. 301–321. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-28577-7_25
5. Buda, J., Bolonyai, F.: An ensemble model using N-grams and statistical features to identify fake news spreaders on Twitter. In: *CLEF (2020)*
6. Chernyaev, A., Spryskov, A., Ivashko, A., Bidulya, Y.: A rumor detection in Russian tweets. In: Karpov, A., Potapova, R. (eds.) *SPECOM 2020*. LNCS (LNAI), vol. 12335, pp. 108–118. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-60276-5_11
7. Cui, L., Lee, D.: CoAID: COVID-19 healthcare misinformation dataset. arXiv preprint [arXiv:2006.00885](https://arxiv.org/abs/2006.00885) (2020)
8. Da San Martino, G. et al.: SemEval-2020 task 11: detection of propaganda techniques in news articles. In: *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pp. 1377–1414 (2020)
9. Devlin, J., et al.: BERT: pre-training of deep bidirectional transformers for language understanding. arXiv preprint [arXiv:1810.04805](https://arxiv.org/abs/1810.04805) (2018)
10. Elhadad, M.K., Li, K.F., Gebali, F.: COVID-19-FAKES: a Twitter (Arabic/English) dataset for detecting misleading information on COVID-19. In: Barolli, L., Li, K.F., Miwa, H. (eds.) *INCoS 2020*. AISC, vol. 1263, pp. 256–268. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-57796-4_25
11. emoji 0.6.0. <https://pypi.org/project/tweet-emoji/>. Accessed 14 Dec 2020
12. g2tmn at Constraint@AAAI2021 - COVID19 fake news detection in English. https://github.com/oldaandozerskaya/covid_news. Accessed 14 Dec 2020

13. Jwa, H., et al.: exBAKE: automatic fake news detection model based on bidirectional encoder representations from transformers (BERT). *Appl. Sci.* **9**19, 4062 (2019)
14. Kar, D. et al.: No rumours please! a multi-indic-lingual approach for COVID fake-tweet detection. arXiv preprint [arXiv:2010.06906](https://arxiv.org/abs/2010.06906) (2020)
15. Kim, D., Graham, T., Wan, Z., Rizoiu, M.-A.: Analysing user identity via time-sensitive semantic edit distance (t-SED): a case study of Russian trolls on Twitter. *J. Comput. Soc. Sci.* **2**(2), 331–351 (2019). <https://doi.org/10.1007/s42001-019-00051-x>
16. Kruspe, A. et al.: Cross-language sentiment analysis of European Twitter messages during the COVID-19 pandemic. arXiv preprint [arXiv:2008.12172](https://arxiv.org/abs/2008.12172) (2020)
17. Kula, S., Choraś, M., Kozik, R.: Application of the BERT-based architecture in fake news detection. In: Herrero, Á., Cambra, C., Urda, D., Sedano, J., Quintián, H., Corchado, E. (eds.) *CISIS 2019. AISC*, vol. 1267, pp. 239–249. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-57805-3_23
18. Kumar, P., Singh, A.: NutCracker at WNUT-2020 Task 2: robustly identifying informative COVID-19 Tweets using ensembling and adversarial training. In: *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*, pp. 404–408 (2020)
19. Liu, Y. et al.: RoBERTa: a robustly optimized BERT pretraining approach. arXiv preprint [arXiv:1907.11692](https://arxiv.org/abs/1907.11692) (2019)
20. Loper, E., Bird, S.: NLTK: the natural language toolkit. In: *Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics*, pp. 63–70 (2002)
21. Loshchilov I., Hutter F.: Fixing weight decay regularization in Adam. arXiv preprint [arXiv:1711.05101](https://arxiv.org/abs/1711.05101) (2017)
22. Mazza, C., et al.: A nationwide survey of psychological distress among Italian people during the COVID-19 pandemic: immediate psychological responses and associated factors. *Int. J. Environ. Res. Public Health* **17**9, 3165 (2020)
23. Mikhalkova, E., et al.: UTMN at SemEval-2020 Task 11: a kitchen solution to automatic propaganda detection. In: *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pp. 1858–1864 (2020)
24. Morio, G., et al.: Hitachi at SemEval-2020 Task 11: an empirical study of pre-trained transformer family for propaganda detection. In: *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pp. 1739–1748 (2020)
25. Moscadelli, A., et al.: Fake news and COVID-19 in Italy: results of a quantitative observational study. *Int. J. Environ. Res. Public Health* **17**16, 5850 (2020)
26. Müller, M., Salathé, M., Kummervold, P.E.: COVID-Twitter-BERT: a natural language processing model to analyse COVID-19 content on Twitter. arXiv preprint [arXiv:2005.07503](https://arxiv.org/abs/2005.07503) (2020)
27. Nguyen, D.Q., et al.: WNUT-2020 Task 2: identification of informative COVID-19 English tweets. In: *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*, pp. 314–318 (2020)
28. Patwa, P., et al.: Fighting an infodemic: COVID-19 fake news dataset. arXiv preprint [arXiv:2011.03327](https://arxiv.org/abs/2011.03327) (2020)
29. Patwa P. et al.: Overview of CONSTRAINT 2021 Shared Tasks: Detecting English COVID-19 Fake News and Hindi Hostile Posts. In: Chakraborty, T., Shu, K., Bernard, R., Liu, H., Akhtar, M.S. (eds.) *Proceedings of the First Workshop on Combating Online Hostile Posts in Regional Languages during Emergency Situation, CONSTRAINT 2021, CCIS*, vol. 1402, pp. 42–53. Springer, Cham (2021)

30. Paszke, A., et al.: PyTorch: an imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems*, pp. 8026–8037 (2019)
31. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011)
32. Peinelt, N., Nguyen, D., Liakata, M. tBERT: topic models and BERT joining forces for semantic similarity detection. In: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 7047–7055 (2020)
33. Pizarro, J.: Using N-grams to detect fake news spreaders on Twitter. In: *CLEF* (2020)
34. Rangel, F., et al.: Overview of the 8th author profiling task at PAN 2020: profiling fake news spreaders on Twitter. In: *CLEF* (2020)
35. Reis, J.C.S., et al.: Supervised learning for fake news detection. *IEEE Intell. Syst.* **234**, 76–81 (2019)
36. Shaar, S., et al.: Overview of CheckThat! 2020 English: automatic identification and verification of claims in social media. arXiv preprint [arXiv:2007.07997](https://arxiv.org/abs/2007.07997) (2020)
37. Shahi, G.K., Nandini, D.: FakeCovid-a multilingual cross-domain fact check news dataset for COVID-19. arXiv preprint [arXiv:2006.11343](https://arxiv.org/abs/2006.11343) (2020)
38. Shu, K., et al.: Fake news detection on social media: a data mining perspective. *ACM SIGKDD Explor. Newsl.* **119**, 22–36 (2017)
39. Tang, L.: UZH at SemEval-2020 task 3: combining BERT with WordNet sense embeddings to predict graded word similarity changes. In: *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pp. 166–170 (2020)
40. Thorne, J., et al.: FEVER: a large-scale dataset for fact extraction and VERification. In: *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, vol. 1(Long Papers), pp. 809–819 (2018)
41. Thorne, J., et al.: The FEVER2.0 shared task. In: *Proceedings of the Second Workshop on Fact Extraction and VERification (FEVER)*, pp. 1–6 (2019)
42. Tran, K.V., et al.: UIT-HSE at WNUT-2020 task 2: exploiting CT-BERT for identifying COVID-19 information on the Twitter social network. In: *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*, pp. 383–387 (2020)
43. tweet-preprocessor 0.6.0. <https://pypi.org/project/tweet-preprocessor/>. Accessed 14 Dec 2020
44. Vijjali, R., et al.: Two stage transformer model for COVID-19 fake news detection and fact checking. arXiv preprint [arXiv:2011.13253](https://arxiv.org/abs/2011.13253) (2020)
45. Williams, E., Rodrigues, P., Novak, V.: Accenture at CheckThat! 2020: if you say so: post-hoc fact-checking of claims using transformer-based models. arXiv preprint [arXiv:2009.02431](https://arxiv.org/abs/2009.02431) (2020)
46. Wolf, T., et al.: Transformers: state-of-the-art natural language processing. In: *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, pp. 38–45 (2020)
47. Wu, S.H., Chien, S.L.: A BERT based two-stage fake news spreaders profiling system. In: *CLEF* (2020)
48. Yang, C., Zhou, X., Zafarani, R.: CHECKED: Chinese COVID-19 fake news dataset. arXiv preprint [arXiv:2010.09029](https://arxiv.org/abs/2010.09029) (2020)

49. Zhang, T., et al.: BDANN: BERT-based domain adaptation neural network for multi-modal fake news detection. In: 2020 International Joint Conference on Neural Networks (IJCNN), pp. 1–8. IEEE (2020)
50. Zhou, X., et al.: Fake news: fundamental theories, detection strategies and challenges. In: Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, pp. 836–837 (2019)