CIVIC-UPM at CheckThat! 2021: Integration of Transformers in Misinformation Detection and Topic Classification

Álvaro Huertas-García[1,2], Javier Huertas-Tato[1], Alejandro Martín[1] and David Camacho[1]

[1]Department of Computer System Engineering, Universidad Politécnica de Madrid, Calle de Alan Turing, 28031, Madrid, Spain

[2]Department of Computer Sciences, Universidad Rey Juan Carlos, Calle Tulipán, 28933, Madrid, Spain

Abstract

Online Social Networks (OSNs) growth enables and amplifies the quick spread of harmful, manipulative and false information that influence public opinion while sow conflict on social or political issues. Therefore, the development of tools to detect malicious actors and to identify low-credibility information and misinformation sources is a new crucial challenge in the ever-evolving field of Artificial Intelligence. The scope of this paper is to present a Natural Language Processing (NLP) approach that uses Doc2Vec and different state-of-the-art transformer-based models for the CLEF2021 Checkthat! lab Task 3. Through this approach, the results show that it is possible to achieve 41.43% macro-average F1-score in the misinformation detection (Task A) and 67.65% macro-average F1-score in the topic classification (Task B).

Keywords

Misinformation, Social Media, Topic Modeling, Fact-checking

1. Introduction

Misleading information spreads on the Internet at an incredible speed and Online Social Networks (OSNs) amplify the quick spread of harmful, manipulative and false information. This phenomenon undermines the integrity of online conversations, influences public opinion, and originates conflicts on social, political, or health issues [1]. In particular, since COVID-19 emerged in Wuhan, China, in December 2019, the public has been bombarded with vast quantities of information, much of which is not checked, leading the World Health Organization (WHO) to coin this situation as the term *infodemic* [1, 2]. Therefore, the development of tools devoted to detecting malicious actors (e.g. bots and trolls) and identifying low-credibility information and misinformation sources is a new crucial challenge. Throughout this paper, we will use the term *misinformation* instead of *fake news* following the recommendations of the

CLEF 2021 – Conference and Labs of the Evaluation Forum, September 21–24, 2021, Bucharest, Romania
alvaro.huertas.garcia@alumnos.upm.es (Á. Huertas-García); javier.huertas.tato@upm.es (J. Huertas-Tato); alejandro.martin@upm.es (A. Martín); david.camacho@upm.es (D. Camacho)
https://github.com/Huertas97 (Á. Huertas-García)
0000-0003-2165-0144 (Á. Huertas-García); 0000-0003-4127-5505 (J. Huertas-Tato); 0000-0002-0800-7632 (A. Martín); 0000-0002-5051-3475 (D. Camacho)

Poynter Institute[1] and the Council of Europe as they consider it inadequate to describe the complexity of the information disorder ecosystem [3].

The scope of this paper is to describe a Natural Language Processing (NLP) approach that makes use of Machine Learning (ML) and Deep Learning (DL) techniques for the CLEF2021 Checkthat! lab Task 3 [4, 5]. In this competition, we carry out a comparative study between the classical Doc2Vec algorithm [6] as document feature extractor combined with ML classifiers, and fine-tuned state-of-the-art models based on Transformers such as T5 [7], RoBERTa [8], Electra [9] and Longformers [6].

This paper is organized into the following sections: Section 2 provides a general view of some related works on misinformation detection and the description of the Checkthat! lab task [4]. Section 3 introduces our proposed approach. Section 4 describes the results from the experiments conducted. Finally, the conclusions are covered in Section 5.

2. Task Description and Related Work

In recent years, there has been growing interest in detecting misinformation [10, 11, 12]. Since 2017, Checkthat! organizers have proposed different tasks of misinformation detection such as automatic identification and verification of claims, check-worthiness, or evidence retrieval [13, 14]. In addition, other authors [12] have committed to combating the misinformation generated during the COVID-19 pandemic by collecting data since the pandemic's outbreak to explore the impact of fact-checkers on misinformation.

The current task addressed in this paper of misinformation detection Checkthat! lab at CLEF 2021 [4, 15, 16] is divided into two subtasks: Task A and Task B. Task A is designed to classify a set of news into four classes (false, partially false, true, other) [17]. On the other hand, Task B consists of classifying a subset of news from Task A into six topical categories: health, economy, crime, climate, elections, and education [18]. Both subtasks share that the text data is divided into the title and the body of news, and that they are a multi-class classification problem with imbalanced data (see Table 1). Therefore, the official evaluation metric is the macro-averaged F1-score. The steps used by the organizers for the data collection were defined in the research presenting the AMUSED framework [19]. It is important to point out that during the data exploration some inconsistencies were found. For example, in Task A, some news titles and bodies seemed to be unrelated. Moreover, in Task B, the title and the body fields appeared to be swapped as the length of the title was longer than the length of the body.

Regarding previous related work in the literature, the appearance of the attention-based method in 2017 [20] paved the way for the development of transformer architectures such as Bidirectional Encoder Representations from transformers (BERT) [21]. Jwa et al. [22] were among the first to develop a model based on BERT for detecting misinformation. The authors conclude that fine-tuning the model in the specific task leads to better results than traditional approaches, such as using a simple classifier model based on TF-IDF and cosine similarity to classify news [23]. Nevertheless, in the literature, there are also examples of using classical techniques such as Doc2Vec [6] to deal with long text documents in tasks related to the fight against misinformation [24].

[1] https://www.poynter.org/

Table 1
Checkthat! lab-CLEF2021 Task 3 breakdown according to its classes.

	Class	Count
Task A	false	465
	partially false	217
	true	142
	other	76
Task B	health	127
	climate	49
	economy	43
	crime	39
	elections	32
	education	28

Unlike Doc2Vec, one of the main transformer-based models' limitations on Natural Language Processing (NLP) tasks is the text length. The average text length in Task A is 4,167 and 286 words in body and title, with a maximum of 32,767 and 9,960 words, respectively. In Task B, in body and title, the average is 4,980 and 566 words, and the maximum is 32,767 and 16,524 words, respectively. Long sequences of text are disproportionately expensive for transformers because attention is quadratic to the sequence length [21]. For this reason, recently, a new method has been proposed, namely Longformer. The authors of Longformer [25] developed a model with an attention mechanism that scales linearly with sequence length by replacing the full self-attention mechanism with the combination of local windowed attention and global attention to have in to account larger interactions without increasing the computation, making it easy to process documents of thousands of tokens. Furthermore, a recent research [26] includes Longformers in a framework for jointly predicting rumor stance and veracity on the dataset released at SemEval 2019 RumorEval [27].

3. Proposed approaches methodology

This section describes the proposed approaches for Tasks A and B of Checkthat! lab CLEF2021. As described in the previous section, the training data for both subtasks contains two text data fields, title and body news. To obtain the best results and avoid overfitting, we reserved 20% of the training data split in a stratified way as a development set. Table 2 summarizes the hyperparameters tuned for both tasks using their respective development set. It is essential to highlight that for each subtask using only titles, only body texts, or title and body texts as data input is explored.

Two remarkable hyperparameters for transformer-based model approaches are the *sliding window* and *oversampling*. As previously mentioned, transformers models typically have a restriction on the maximum length allowed for a sequence. A plausible strategy to overcome this limitation is to use the sliding window approach introduced by Zhang et al. [28]. Here, any sequence exceeding the maximum length is split into several windows (sub-sequences), and each one is assigned the label from the original sequence. We explored the use of this technique, and to minimize any information loss that hard cutoffs between two windows may cause, we applied 20% of overlapping between the sub-sequences. Finally, we explored to over-sample the unbalanced data so that all classes had the same frequency as the most abundant class using the RandomOversampler from *imblearn*[2] package.

Table 2
Hyperparameters optimized during the development of the proposed approaches. C is the inverse of regularization strength, logspace is the logarithmic sequence (start base, end base, number of elements).

Optimization	Method	Hyperparameters	Values
Transformer-based model	Grid and Bayesian Search	learning rate	min = 1e-6 , max = 1e-3
		epochs	min = 1, max = 20
		weight decay	min = 0 , max = 1
		gradient accumulation steps	min = 1 , max = 4
		scheduler	constant_schedule_with_warmup cosine_schedule_with_warmup polynomial_decay_schedule_with_warmup
		sliding window	True, False
		over-sampling	True, False
Doc2Vec	Grid Search	dimensions	[50, 75, 100]
		window size	[15, 20, 30]
		min count	[2, 5]
RF	Grid search with CV = 5	n_estimators	[5, 10, 15, 30]
		max_depth	[3, 5, 10, 15, 20]
		min_samples_split	[2, 5, 10]
		min_samples_leaf	[1, 2, 4]
		max_features	[2, 3, "auto"]
		min_samples_split	[8, 10, 12]
LR1	Grid search with CV = 5	C	logspace(-3, 2, 8)
LR2	Grid search with CV = 5	C	logspace(-3, 2, 8)
Elastic Net	Grid search with CV = 5	C	logspace(-3, 2, 8)
		L1_ratio	[0 , 0.33333333, 0.66666667, 1]
SVC	Grid search with CV = 5	C	numpy logspace(-3, 2, 10)
		Kernel	polynomial, RBF, linear
		Gamma	logspace(-3, 3, 10)

3.1. Task A

To carry out the Task A, two approaches are tested. The first one is based on the use of the classical Doc2Vec algorithm [6] as document feature extractor combined with Machine Learning (ML) classifiers. The second approach takes advantage of different state-of-the-art transformer-based models [20, 21] to extract dense embeddings with a linear layer on top to classify the documents into four categories.

3.1.1. Doc2Vec approach

Doc2Vec represents documents into dense vectors named document or paragraph embeddings. This algorithm extends the idea of Word2Vec [29, 30], adding a new paragraph representation that is trained along word embeddings to develop document-level embeddings so that documents of differing lengths can be represented by fixed-length vectors [6]. These dense document vectors can be obtained by concatenating the paragraph vector with the word vectors to predict a target

[2] https://github.com/scikit-learn-contrib/imbalanced-learn

word, or predicting sample words from the paragraph using the paragraph vector. These two implementations of Doc2Vec are named PD-DM and PD-DBOW, respectively. The Doc2Vec models are obtained from Gensim library [31]. We explore the use of PD-DM, PD-DBOW, and the combination of both models as feature extractors for this classification task.

The classifiers tested were Naive Bayes (NB), Random Forest (RF), Logistic Regression with L1 and L2 regularization (LR1 and LR2, respectively), Elastic Net, and Support Vector Classifier (SVC).

The data processing for this approach consists of different steps. The *ftfy* package [32] is used to repair Unicode and emoji errors, and the *ekphrasis* package [33] for lower-casing, normalizing percentages, time, dates, emails, phones and numbers. Abbreviations are expanded using *contractions* package[3] and word tokenization, stop-word removal, punctuation removal, and word lemmatization is carried out using the NLTK toolkit [34].

3.1.2. Transformers approach

In this approach, we use different transformer-based models to classify the Task A news. The models tested were T5 small and T5 base [7], Longformer base [6], RoBERTa base [8] and DistilRoBERTa base [8, 35]. The data processing procedure for this approach consists of repairing Unicode and emoji errors with *ftfy* package [32] and normalizing emails, phones and URLs with *ekphrasis* package [33].

Finally, the model with the best performance on the development set is selected to boost its performance by incorporating more data from related tasks: Kaggle's KDD2020[4] and Clickbait news detection[5] competitions. KDD2020 competition consists of distinguishing fake claims from authentic ones. On the other hand, Clickbait detection is focused on classifying articles into news, clickbait, and other.

3.2. Task B

The proposed approach for Task B is based on transformer-based models. The models tested were: Electra base [9], T5 base [7], RoBERTa base [8] and DistilRoBERTa base [8, 35]. As for the transformer-based model approach for Task A, the data processing procedure consists of repairing Unicode and emoji errors with *ftfy* package [32] and normalizing emails, phones and URLs with *ekphrasis* package [33].

In addition, multi-task training was explored in the case of the T5 base model. The model was trained on Task B and Kaggle's Ag News task[6]. Ag News is a topic classification competition with 120k news grouped into 4 categories: World, Sports, Business, and Sci-Tech.

[3] https://github.com/kootenpv/contractions
[4] https://www.kaggle.com/c/fakenewskdd2020/overview
[5] https://www.kaggle.com/c/clickbait-news-detection
[6] https://www.kaggle.com/amananandrai/ag-news-classification-dataset

Table 3
Performance of Doc2Vec models in Checkthat! lab CLEF2021 development set. Performance is reported as macro-averaged F1-score×100.

Model	Data input	Classifier	F1-score
PV-DM	Title	RF	25.81
PV-DBOW	Title	LR2	23.67
PV-DM	Body	LR2	22.74
PV-DBOW	Body	NB	27.72
PV-DM & PV-DBOW	Title	LR2	29.23
PV-DM & PV-DBOW	Body	NB	24.96
PV-DM & PV-DBOW	Title & Body	LR2	25.93

Table 4
Performance of Transformer-based models in Checkthat! lab CLEF2021 Task A development set. Performance is reported as macro-averaged F1- score×100.

Model	Train data	Data input	Oversampling	Sliding window	F1-score
T5 small	Checkthat!	Title	True	False	34.52
T5 base	Checkthat!	Title & Body	False	True	37.11
Longformer base	Checkthat!	Body	False	True	45.59
RoBERTa base	Checkthat!	Body	True	True	48.62
DistilRoBERTa base	**Checkthat!**	**Body**	**True**	**True**	**50.96**
DistilRoBERTa base	Checkthat! & KD2020	Body	True	True	42.17
DistilRoBERTa base	Checkthat! & Clickbait	Title	True	True	35.37

4. Experiments and Results

4.1. Task A

Table 3 reports the performance of Doc2Vec models evaluated in the development set. The best macro F1-score (29.23%) is achieved using the title field as input data and combining features from PV-DM and PV-DBOW models with Logistic Regression classifier with L2 regularization. Remarkably, this same approach worsens when the input data includes the body text field: 24.96% F1-score only with body text and 25.93% F1-score with title and body texts.

Regarding the transformer-based model approach, Table 4 details the performance of the models, the training data, the type of data input, and if oversampling and sliding window techniques are used during training.

As expected, our experiments show that state-of-the-art transformer-based models outperform the classical Doc2Vec algorithms. The best performance, 50.96% macro-averaged F1-score, is achieved with DistilRoBERTa base, a distilled version of RoBERTa base, using the body field from Checkthat! data as data input with oversampling and sliding window for dealing with long texts. The hyperparameters selected for this model were polynomial decay scheduler with warmup, one step for gradient accumulation, 0.04731 as weight decay, and learning rate equals to 9.468e-5. Significantly, the performance of the model obtained using the same hyperparameters without oversampling and without sliding window was 39.61% macro-averaged

Table 5
Performance of transformer-based models in Checkthat! lab CLEF2021 Task B development set. Performance is reported as macro-averaged F1- score×100.

Model	Train Data	Data Input	Oversampling	Sliding Window	F1-score
Electra base	Checkthat!	Body	True	True	87.08
T5 base	Checkthat!	Title and Body	True	False	82.49
T5 base multitasking	Checkthat! and Ag News	Title	True	False	33.87
DistilRoBERTa base	Checkthat!	Body	True	False	88.82
RoBERTa base	**Checkthat!**	**Title and Body**	**True**	**False**	**91.22**

F1-score. Remarkably, the introduction of new related data from KDD2020 and Clickbait news detection competitions did not improve the performance on Checkthat! lab Task A. Moroever, the Clickbait task has a more noticeable impact on performance, suggesting that less related tasks have more impact on performance.

The official test results for the best model on the development set, DistilRoBERTa base, are shown in Table 6.

4.2. Task B

Table 5 compares the performance of the different transformer-based models evaluated on Checkthat! Task B development set.

All the models get the best performance applying oversampling. Unlike Task A, the sliding window technique does not seem crucial for the topic classification. The reason could be due in part to the different complexity of the tasks. The classification of news articles according to the veracity of their information requires a high level of text comprehension. On the other hand, the topic analysis may not require such a high level of comprehension, and using the sliding window technique could be not as important.

Turning now to the T5 base model multi-task training, we can note from Table 5 that multi-task training on Ag News and Checkthat! lab Task B reduces the performance from 82.49% to 33.87% of the T5 base model on Task B compared with the T5 base model exclusively fine-tuned for Task B.

Finally, the best macro-average F1-score (91.22%) is obtained by RoBERTa base model trained using the title and body text fields from Checkthat! lab Task B data with oversampling and without sliding window. The hyperparameters used for training this model were constant schedule with warmup, one step for gradient accumulation, one as weight decay, and the learning rate equals 5.583e-5. This model was submitted to the competition, and the official test result is shown in Table 6.

5. Conclusion

In this work, we have proposed a NLP approach for misinformation detection Task A and topic classification Task B from the CLEF2021 Checkthat! lab Task 3 [4, 15]. Our work has led us to conclude that transformer-based models fine-tuned explicitly for the tasks have achieved the best performance. In Task A, the results indicate that the transformer-based models

Table 6
Official results of the subtasks from CLEF2021 Checkthat! lab Task 3. Performance is reported as macro-averaged F1- score×100.

Model	Task	Training Data	F1-score
DistilRoBERTa base	A	Checkthat!	41.43
RoBERTa base	B	Checkthat!	67.65

outperform the classical Doc2Vec model. Oversampling proves to be a valuable technique to deal with unbalanced data in both tasks. However, the sliding window technique to overcome the maximum length transformers' limitation shows different effects in Task A and Task B. Finally, we achieved a macro-average F1-score of 41.43% in Task A and 67.65% in Task B. In future work, we will most likely test new architectures, such as Hierarchical Attention Networks, and add more related data to boost the transformer-based model performance.

Acknowledgements

This work has been partially supported by the following grants and funding agencies: Spanish Ministry of Science and Innovation under TIN2017-85727-C4-3-P (DeepBio) grant, by Comunidad Autónoma de Madrid under S2018/TCS-4566 grant (CYNAMON), and by BBVA FOUNDATION GRANTS FOR SCIENTIFIC RESEARCH TEAMS SARS-CoV-2 and COVID-19 under the grant: "*CIVIC: Intelligent characterisation of the veracity of the information related to COVID-19*". Relevant parts of this research is a result of the project IBERIFIER - Iberian Digital Media Research and Fact-Checking Hub, funded by the European Commission under the call CEF-TC-2020-2 (European Digital Media Observatory), grant number 2020-EU-IA-0252. Finally, the work has been supported by the Comunidad Autónoma de Madrid under Convenio Plurianual with the Universidad Politécnica de Madrid in the actuation line of "*Programa de Excelencia para el Profesorado Universitario*".

References

[1] S. B. Naeem, R. Bhatti, The Covid-19 'infodemic': a new front for information professionals, Health Information & Libraries Journal 37 (2020) 233–239. doi:10.1111/hir.12311.

[2] M. Cinelli, W. Quattrociocchi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, A. Scala, The COVID-19 social media infodemic, Scientific Reports 10 (2020) 16598. doi:10.1038/s41598-020-73510-5.

[3] A. Estrada-Cuzcano, K. Alfaro-Mendives, V. Saavedra-Vásquez, Disinformation y misinformation, posverdad y fake news: precisiones conceptuales, diferencias, similitudes y yuxtaposiciones, Información, cultura y sociedad (2020) 93–106. doi:10.34096/ics.i42.7427.

[4] P. Nakov, G. Da San Martino, T. Elsayed, A. Barrón-Cedeño, R. Míguez, S. Shaar, F. Alam, F. Haouari, M. Hasanain, N. Babulkov, A. Nikolov, G. K. Shahi, J. M. Struß, T. Mandl, The CLEF-2021 CheckThat! lab on detecting check-worthy claims, previously fact-checked

claims, and fake news, in: Proceedings of the 43rd European Conference on Information Retrieval, ECIR '21, Lucca, Italy, 2021, pp. 639–649. URL: https://link.springer.com/chapter/10.1007/978-3-030-72240-1_75.
[5] G. K. Shahi, J. M. Struß, T. Mandl, Overview of the CLEF-2021 CheckThat! lab task 3 on fake news detection, in: Working Notes of CLEF 2021—Conference and Labs of the Evaluation Forum, CLEF '2021, Bucharest, Romania (online), 2021.
[6] Q. V. Le, T. Mikolov, Distributed representations of sentences and documents, 2014. arXiv:1405.4053.
[7] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, P. J. Liu, Exploring the limits of transfer learning with a unified text-to-text transformer, 2020. arXiv:1910.10683.
[8] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov, Roberta: A robustly optimized bert pretraining approach, 2019. arXiv:1907.11692.
[9] K. Clark, M.-T. Luong, Q. V. Le, C. D. Manning, Electra: Pre-training text encoders as discriminators rather than generators, 2020. arXiv:2003.10555.
[10] H. Jwa, D. Oh, K. Park, J. M. Kang, H. Lim, exbake: Automatic fake news detection model based on bidirectional encoder representations from transformers (bert), Applied sciences 9 (2019) 4062.
[11] R. K. Kaliyar, A. Goswami, P. Narang, Fakebert: Fake news detection in social media with a bert-based deep learning approach, Multimedia tools and applications 80 (2021) 11765–11788.
[12] G. Burel, T. Farrell, M. Mensio, P. Khare, H. Alani, Co-spread of misinformation and fact-checking content during the covid-19 pandemic, in: Proceedings of the 12th International Social Informatics Conference (SocInfo), LNCS, 2020. URL: http://oro.open.ac.uk/71786/.
[13] T. Elsayed, P. Nakov, A. Barrón-Cedeño, M. Hasanain, R. Suwaileh, G. Da San Martino, P. Atanasova, Overview of the clef-2019 checkthat! lab: Automatic identification and verification of claims, in: Experimental IR Meets Multilinguality, Multimodality, and Interaction, volume 11696 of *Lecture Notes in Computer Science*, Springer International Publishing, Cham, 2019, pp. 301–321.
[14] A. Barrón-Cedeño, T. Elsayed, P. Nakov, G. Da San Martino, M. Hasanain, R. Suwaileh, F. Haouari, N. Babulkov, B. Hamdan, A. Nikolov, S. Shaar, Z. S. Ali, Overview of checkthat! 2020: Automatic identification and verification of claims in social media, in: Experimental IR Meets Multilinguality, Multimodality, and Interaction, Lecture Notes in Computer Science, Springer International Publishing, Cham, 2020, pp. 215–236.
[15] P. Nakov, G. Da San Martino, T. Elsayed, A. Barrón-Cedeño, R. Míguez, S. Shaar, F. Alam, F. Haouari, M. Hasanain, N. Babulkov, A. Nikolov, G. K. Shahi, J. M. Struß, T. Mandl, S. Modha, M. Kutlu, Y. S. Kartal, Overview of the CLEF-2021 CheckThat! Lab on Detecting Check-Worthy Claims, Previously Fact-Checked Claims, and Fake News, in: Proceedings of the 12th International Conference of the CLEF Association: Information Access Evaluation Meets Multiliguality, Multimodality, and Visualization, CLEF '2021, Bucharest, Romania (online), 2021.
[16] G. K. Shahi, D. Nandini, FakeCovid – a multilingual cross-domain fact check news dataset for covid-19, in: Workshop Proceedings of the 14th International AAAI Conference on Web and Social Media, 2020. URL: http://workshop-proceedings.icwsm.org/pdf/2020_14.pdf.

[17] G. K. Shahi, A. Dirkson, T. A. Majchrzak, An exploratory study of covid-19 misinformation on twitter, Online Social Networks and Media 22 (2021) 100104.
[18] G. K. Shahi, A multilingual domain identification using fact-checked articles: A case study on covid-19 misinformation, arXiv preprint (2021).
[19] G. K. Shahi, Amused: An annotation framework of multi-modal social media data, arXiv preprint arXiv:2010.00502 (2020).
[20] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, Attention is all you need, 2017. arXiv:1706.03762.
[21] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, Bert: Pre-training of deep bidirectional transformers for language understanding, arXiv preprint arXiv:1810.04805 (2018).
[22] H. Jwa, D. Oh, K. Park, J. M. Kang, H. Lim, exbake: Automatic fake news detection model based on bidirectional encoder representations from transformers (bert), Applied Sciences 9 (2019). doi:10.3390/app9194062.
[23] B. Riedel, I. Augenstein, G. P. Spithourakis, S. Riedel, A simple but tough-to-beat baseline for the fake news challenge stance detection task, 2018. arXiv:1707.03264.
[24] B. Anjali, R. Reshma, V. Geetha Lekshmy, Detection of counterfeit news using machine learning, in: 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), volume 1, 2019, pp. 1382–1386. doi:10.1109/ICICICT46008.2019.8993330.
[25] I. Beltagy, M. E. Peters, A. Cohan, Longformer: The long-document transformer, 2020. arXiv:2004.05150.
[26] A. Khandelwal, Fine-tune longformer for jointly predicting rumor stance and veracity (2020).
[27] G. Gorrell, K. Bontcheva, L. Derczynski, E. Kochkina, M. Liakata, A. Zubiaga, Rumoureval 2019: Determining rumour veracity and support for rumours, 2018. arXiv:1809.06683.
[28] Z. Wang, P. Ng, X. Ma, R. Nallapati, B. Xiang, Multi-passage bert: A globally normalized bert model for open-domain question answering, 2019. arXiv:1908.08167.
[29] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, J. Dean, Distributed representations of words and phrases and their compositionality, 2013. arXiv:1310.4546.
[30] T. Mikolov, K. Chen, G. Corrado, J. Dean, Efficient estimation of word representations in vector space, 2013. arXiv:1301.3781.
[31] R. Řehůřek, P. Sojka, Software Framework for Topic Modelling with Large Corpora, in: Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks, ELRA, Valletta, Malta, 2010, pp. 45–50.
[32] R. Speer, ftfy, Zenodo, 2019. doi:10.5281/zenodo.2591652, version 5.5.
[33] C. Baziotis, N. Pelekis, C. Doulkeridis, Datastories at semeval-2017 task 4: Deep lstm with attention for message-level and topic-based sentiment analysis, in: Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 747–754.
[34] E. Loper, S. Bird, 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 - Volume 1, ETMTNLP '02, Association for Computational Linguistics, USA, 2002, p. 63–70. URL: https://doi.org/10.3115/1118108.1118117. doi:10.3115/1118108.1118117.

[35] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. L. Scao, S. Gugger, M. Drame, Q. Lhoest, A. M. Rush, Huggingface's transformers: State-of-the-art natural language processing, 2020. arXiv:1910.03771.