IJICTR
International Journal of Information & Communication Technology Research

Volume 14- Number 4 – 2022 (46 -54)

Stance Detection Dataset for Persian Tweets

Mohammad Hadi Bokaei
ICT Research Institute (ITRC)
Tehran, Iran
mh.bokaei@itrc.ac.ir

Mojgan Farhoodi*
ICT Research Institute (ITRC)
Tehran, Iran
farhoodi@itrc.ac.ir

Mona Davoudi Shamsi
ICT Research Institute (ITRC)
Tehran, Iran
davoudi@itrc.ac.ir

Received: 3 March 2022 – Revised: 11 April 2022 - Accepted: 26 June 2022

Abstract—**Stance detection aims to identify an author's stance towards a specific topic which has become a critical component in applications such as fake news detection, claim validation, author profiling, etc. However, while the stance is easily detected by humans, machine learning models are falling short of this task. In the English language, due to having large and appropriate e datasets, relatively good accuracy has been achieved in this field, but in the Persian language, due to the lack of data, we have not made significant progress in stance detection. So, in this paper, we present a stance detection dataset that contains 3813 labeled tweets. We provide a detailed description of the newly created dataset and develop deep learning models on it. Our best model achieves a macro-average F1-score of 58%. Moreover, our dataset can facilitate research in some fields in Persian such as cross-lingual stance detection, author profiling, etc.**

Keywords: Stance Detection; Fake News; Social Media; Twitter; Persian Dataset; Author Profiling

Article type: Research Article

[DOI: 10.52547/itrc.14.4.46]

I. INTRODUCTION

Social media is a double-edged sword for consuming news. Minimal effort, easy access, and rapid data dissemination on the Internet and social media are increasingly encouraging people to switch from traditional news sources to online ones. Sources such as Twitter, Facebook online news sites, other social media platforms, and the personal blogs of self-proclaimed journalists have become important players in providing news content [1]. So, governments, journalists, and social media platforms are working hard to distinguish authentic news from fake news. The goal of the Fake News Challenge [2] is to explore how artificial intelligence technologies, particularly machine learning and natural language processing, might be leveraged to combat the fake news problem. This process can be divided into several stages [3]. The first useful step in identifying fake news is to find out what other news sources, posts or comments have to say about it. This is why the Fake News Challenge initially focuses on stance detection. In other words, stance detection is the first and most important step in detecting fake news [2], which is still in the early stages of research.

Stance detection is the task of automatically determining from the text whether the author of the text agrees, disagrees, or is neutral towards a proposition or target and it has become a key component in applications such as fake news detection, claim validation, or argument search [5]. The target may be a person, an organization, a government policy, a movement, a product, and so on [4].

In this paper, we present a Persian dataset for reply-to-post stance detection which contains 3813 tweets. This corpus has been used in the development of an automated stance detection system based on transformers.

The rest of the paper is organized as follows: Section 2 reviews the previous works on the current study. Section 3 discusses our methodology and corpus information. In section 4, the experimental reports are

* Corresponding Author

presented, including evaluation metrics, error analysis, and results. Section 5 concludes the survey and suggests future directions in this area.

TABLE I. COMPARISON OF STANCE DETECTION DATASETS

Reference	Source	Type	Language	Size
[6]	Twitter	Target-specific	English	4,870
[7]	News articles	Claim-based	English	2,595
[8]	Twitter	Multi-target	English	4,455
[9]	Twitter	Claim-based	English	5,568
[10]	Twitter	Target-specific	English	3,545
[11]	Twitter, Reddit	Claim-based	English	8,574
[12]	Twitter	Target-specific	English	51,284
[13]	Twitter	Target-specific	English	21,574
[14]	Web sites	Claim-based	Persian	2.029 (Headline-Claim)
				1997 (Article-Claim)
Our dataset	Twitter	Claim-based	Persian	7.738

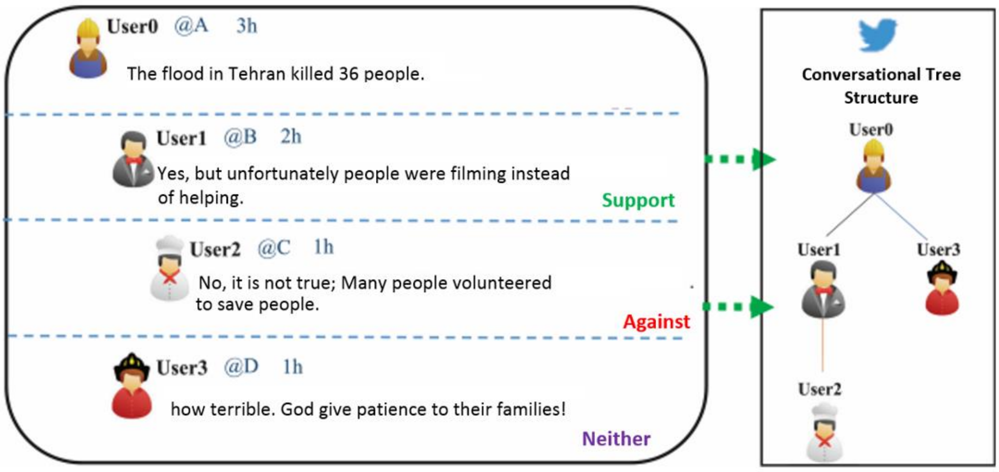

Figure 1. Tree Structure of Social Media Conversational

II. RELATED WORKS

Stance detection (also known as stance classification, stance prediction, and stance analysis) is a problem related to social media analysis, natural language processing, and information retrieval that aims to define the position of a person from the text they produce, toward a target (a concept, idea, event, etc.) that is explicitly stated in the text, or only implied [30].

The stance detection task needs the presence of a defined target to detect the stance toward it. In the existing literature, stance detection can be categorized into different types: target-specific, multi-target, and claim-based. The majority of methods of stance detection are target-specific stance detection which aims at detecting the stance expressed in the text towards a specific target [15]. Many studies have focused on this issue [16, 17, 10]. For example, [16] provided a set of resources on topics related to politics and then applied various features based on the textual content of the tweet and different features based on contextual information for the stance detection task. In [10], a novel dataset was created which includes up to 3000 English-Hindi tweets with opinions toward Demonetization that was implemented in India in 2016. More recently, since people often comment on multiple target entities in the same text, multi-target stance detection was designed. The goal of multi-target stance detection is to jointly learn the social media users' orientation toward two or more targets for a single topic [8, 18]. Claim-based stance detection is considered a suitable method to analyze the veracity of the news. For that reason, claim-based stance detection has been heavily used for rumor resolution studies [9, 11, 20].

On the other hand, researches show that various approaches have been used for the task of stance detection, such as traditional machine learning approaches like SVM[2] and logistic regression [31, 32],

[2] Support Vector Machine

[DOI: 10.52547/itrc.14.4.46]

deep learning approaches such as LSTM[3] and CNN[4] [33, 34], and ensemble methods [10, 35]. Although machine learning approaches like SVM are the most commonly used method until 2019 [30], recent studies tend to apply deep learning algorithms [38, 39].

TABLE II. SOME EXAMPLES OF OUR DATASET

Main Post	Reply Post	Stance
Ahwaz steel workers went on strike for the third day and gathered in front of the Khuzestan governor's office, protesting against low wages and job insecurity.	We need fundamental change, the lack of "job security" is the most important concern of workers.	Support
The US space agency, NASA, has released a report on the effects of global warming that some parts of the world, including Iran, will not be viable for another 30 years.	Raise your life expectancy. We are going to see happy days. For drought, it can be solved if the work is left to the skilled	Against
During the years 1393 to 1400, on average, about 4% of electricity consumption was saved due to changes in the working hours of offices, which is equivalent to 106 billion kilowatt hours over 7 years.	Excuse me, do you have a scientific reference for this issue?	Neither

In most of these studies, the available datasets for stance detection focus on English texts. [21] was the first study of automatic stance classification, which propose a semantic model for predicting claim stance based on a dataset that includes 55 topics. In this work, they got IBM argumentative structure dataset [37] that contains claims and evidence for 33 controversial topics and developed it into 55 topics. Topics were randomly selected from the debate motions database at the International Debate Education Association (IDEA)[5].

Similar to it, [22] performs rumor stance classification by using the dataset created by [21]. RumourEval which is a shared task, organized as part of SemEval in 2017 and 2019, hosts a dataset of annotated English tweets and their stance (favor of or against) towards the various targets of interest commonly known and debated in the United States, such as ‘Atheism’, ‘Climate Change’, ‘Feminism’, ‘Hillary Clinton’, and ‘Abortion’ and so on [6, 11]. This dataset includes 4780 samples that its annotations were performed by crowdsourcing and then several techniques were employed to encourage high-quality annotations. Of course, some studies have been done in other languages as well, for example, there are available datasets of tweets in French [16], Italian [23], Russian [24], and Catalan and Spanish [17]. But [14] is the only dataset related to the Persian language. It collects claims from Fakenews[6] and Shayeaat[7] websites. After collecting the articles, it assigned three labels to each claim. The first label is the stance of the article toward the claim. The second label is the stance of the article's headline based according to the claim and the third one is the stance of the article on its headline.

Due to the scarcity of the Persian datasets in this field, we started to build a dataset, which the details are given in section 2. [13] compares some existing English stance detection datasets and we also added two Persian data sets to it. Table 1 shows a summary of this information.

We can observe that the language of the existing stance detection dataset is English except for the dataset created by [14]. However, this work collected real rumors from the websites and then it found similar news to them to assign stance of this news to collected rumors; but we used people's responses to the same claims to find the stance. Also, the main difference between our work and recent works is that in recent studies, the main goal is to detect the stance of news or posts concerning the claim to determine whether the related rumor or claim is fake or not; But in our work, determining whether the claim is fake or fact is not the main goal, but we want to know whether the authors of post replies agrees with the author of the claim or not. Therefore, our dataset is also used for account profiling.

III. BUILDING THE DATASET

In this section, we detail the creation and the particularities of our stance detection dataset composed of 3813 tweets collected from Twitter.

A. Data Collection

We used the following two methods to collect data from Twitter:

- We identified the most popular Twitter accounts (according to their follower rate) and then tried to select the original tweets of these accounts randomly from the time 26 March 2018 until 10 August 2018. Then, for each selected tweet, we extracted a maximum of 15 reply tweets. Thus, we collected 2242 pairs of original and reply tweets.

  As shown in Figure 1, conversations on a social network such as Twitter are tree-structured. In other words, the replies are often nested and are triggered by a source tweet that initiated the conversation. What we considered in collecting this dataset are

[3] Long Short-Term Memory
[4] Convolutional Neural Network
[5] http://idebate.org
[6] Fakenews.ir
[7] Shayeaat.ir

[DOI: 10.52547/itrc.14.4.46]

first-level replies, that is, we selected only replies that responded to the source tweet.

- First, by referring to the Factnameh[8], up to 50 news whose authenticity was denied were selected. Then on Twitter, we found tweets similar to this news and used the procedure mentioned in the first method for each of these tweets. Thus, various pairs of original and reply tweets were collected.

Finally, using the above two methods, we prepared a dataset with 8461 samples.

B. Preprocessing

To increase the quality of this dataset, we performed several pre-processing steps:

- We removed tweets with less than 20, or more than 140 characters; because according to our observations, these tweets were usually not informative, in other words, they were either too noisy or contained repetitive phrases.
- We removed duplicate tweets because it reduces the performance of the created model.
- We only kept the tweets in Persian because our goal in this work is to build a Persian stance detection dataset.
- We kept only text tweets without any media such as images or videos, as our goal was to detect the stance through textual content.
- We also removed tweets that contained insulting words.
- We also deleted tweets containing URLs.

Finally, using the mentioned preprocessing steps, the size of the dataset reached 7738 samples.

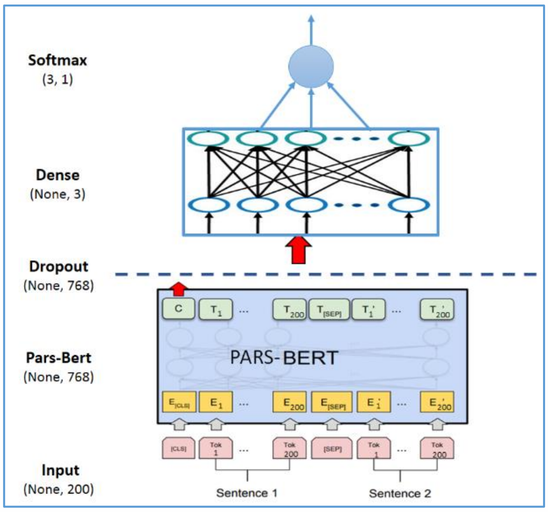

Figure 2. the Schematic of Our model.

C. Data Annotation

The main purpose of this paper is to determine whether the author of the post and the author of the reply post agree on the subject of the main post or not. In other words, whether the stance of both people toward to the issue raised in the post is the same or not. so, we define three followings annotate:

- Support: The stance of the author of the main post and the author of the reply post to the subject of the post is the same.
- Against: The stance of the author of the main post and the author of the reply post to the subject of the post are not the same.
- Neither: It is not possible to detect whether the stance of the two people is the same or not.

We provide some examples of our dataset in Table 2.

To annotate the dataset, we chose two people to annotate the tweets. The data is annotated by both people and if there is no agreement between them, it will be given to the third person to determine the final label. The distribution of stance classes is illustrated in Table 3.

TABLE III. DISTRIBUTION OF STANCE CLASSES

	Support	Against	Neither
Count	3920	2782	1036
Percent	%51	%36	%13

This dataset was labeled by three persons and the inter-annotator agreement is 0.61. we used Cohen's Kappa to calculate the agreement between the annotators.

D. Data Validation

We used a team of 9 people for data labeling. All annotators are native Persian speakers. We prepared a guideline for stance labeling which consists of guides, tips, and various examples. We gave each sample to two people for labeling, and if these two people disagreed, the corresponding sample was given to the third person, and finally, if three annotators assign three different labels, the sample was given to the fourth person for labeling and finally, the majority voting was calculated and considered as the final label.

E. Comparison

Finally, we compare our dataset with the only Persian stance detection in Table 4:

TABLE IV. COMPARISON OF OUR DATASET WITH ANOTHER PERSIAN STANCE DETECTION DATASET

Dataset	Size	Labels	Number	Present
[14]	2029 (Headline-Claim)	Agree	405	20
		Disagree	164	8
		Discuss	802	40
		Unrelated	658	32
	1997	Agree	137	7
		Disagree	206	10

8 https://factnameh.com

[DOI: 10.52547/itrc.14.4.46]

	(Article-Claim)	Discuss	1068	53
		Unrelated	586	30
Our dataset	7738	Support	3920	51
		Against	2782	36
		Neither	1036	13

IV. EXPERIMENTS

This section provides results for evaluating the dataset. In this section, we first talk about the model and experimental setting, then the results are reported. Finally, the errors of the model are discussed to open the way for future studies of stance detection.

A. Model

We experiment with a transformer-based [25] model for our tasks, called Pars-Bert [26]. In 2018, [27] introduced a transformer-based machine learning technique for natural language processing (NLP) pre-training, which stands for Bidirectional Encoder Representations from Transformers (BERT).

Pars-Bert model is a monolingual language model for Persian language with the same configurations as Bert [27], pre-trained on different texts such as news, novels, scientific documents etc. We used its base model and fine-tuned this model using the stance detection corpus. It is followed by a fully-connected network to map the Pars-Bert's outputs to the tag space. The schematic of our model is presented in Fig. 2.

B. Experimental Setting

The number of samples in the dataset is 7738. We used 60% of the corpus as training data, 20% as validation data, and 20% as test data. The learning rate is set to 5e-05 and the batch size and the number of epochs are 32 and 10 respectively. Adam [28] was applied for optimizing the model. We used the TensorFlow library [29] to implement this model. The hyper-parameters have been tuned by evaluation on the validation set to get the highest F1-score. We applied a dropout of rate 0.1 and we used a softmax layer is used as the output layer to create distribution over target labels. Moreover, we used cross entropy as the loss function.

C. Evaluation

- Evaluation Metrics

For evaluating our model, we use accuracy, precision, recall, and F1-score as evaluation metrics.

Precision is the fraction of relevant instances among the retrieved instances, while recall is the fraction of relevant instances that were retrieved. F-measure provides a single score that balances both the concerns of precision and recall in one number and finally, Accuracy is the fraction of predictions our model got right.

- Results

The results of the experiments on our model are reported in Table 5.

TABLE V. EVALUATION CRITERIA ON TEST-DATA

Precision	Recall	F-Measure	Accuracy
0.65	0.64	0.64	0.64

For a more detailed analysis, first, we present the Confusion Matrix for each label in test data with 660 samples (Table 6 to Table 8). The evaluation metrics are calculated for each of them in Fig. 2.

TABLE VI. CONFUSION MATRIX FOR ALL LABELS

	Support	Against	Neither
Support	572	161	52
Against	172	360	41
Neither	61	64	65

TABLE VII. CONFUSION MATRIX FOR "SUPPORT" LABEL

	Support	Others
Support	572	213
Others	233	532

TABLE VIII. CONFUSION MATRIX FOR "AGAINST" LABEL

	Against	Others
Against	360	213
Others	255	750

TABLE IX. CONFUSION MATRIX FOR "NEITHER" LABEL

	Neither	Others
Neither	65	128
Others	83	1277

- Analysis

According to the results represented in Figure 3, our baseline model performs best in the "Neither" stance with an accuracy 82% but in the "Support" and "Against" stance, the accuracy of the algorithm is lower. For a more detailed analysis, we checked the samples in the test data and the results of the algorithm for some of them. In the following, we state the most important reasons for the algorithm's error:

- In general, the stance detection task is very complex and as mentioned in the previous section, the agreement rate of our human annotators was also 61%. In other words, in some cases, the written post or reply is very ambiguous and its understanding is a complicated task even for humans and therefore it is difficult to recognize its stance even by the human tagger. Table 9 shows this problem with two samples in the dataset.

[DOI: 10.52547/itrc.14.4.46]

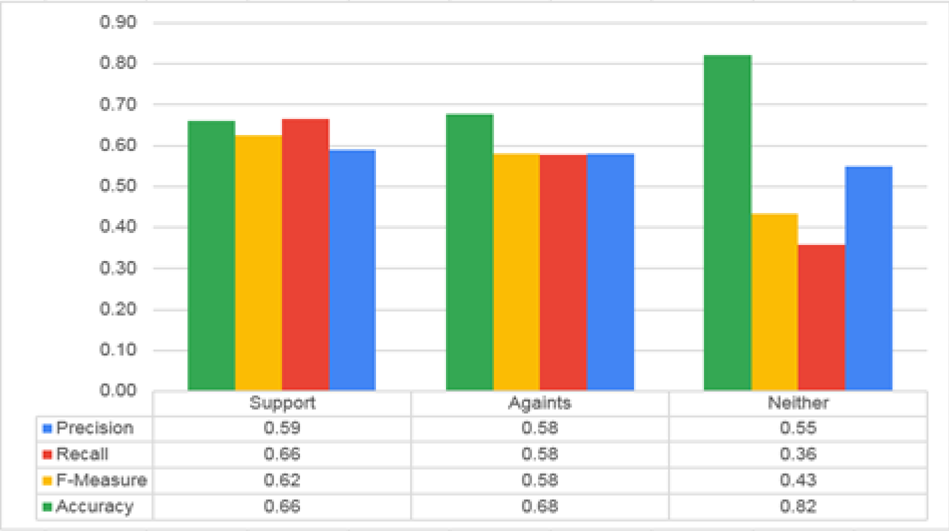

Figure 3. Evaluation Metrics for Each label in Test Data.

TABLE X. SAMPLES OF DATASET

Main Post	Reply Post
The publication of the news of paying the tribute of Levan Dzhagaryan, the Russian ambassador in Tehran, to the monument of Alexander Griboyedov, the initiator of the Treaty of Turkmenchay, has angered the Russian embassy in Tehran.	Yes, the old fox does not stop being cunning!
Real Label: Support	*Predicted Label: Against*
"NASA says: Iran will not be livable in thirty years! The NASA Astronautically Organization announced by publishing a report about the consequences of global warming that some regions of the world, including Iran, will not be habitable until 30 years from 2050. Other countries from the Persian Gulf such as Oman and Kuwait are also on this list.	Does it mean that others live in another Korea?
Real Label: Neither	*Predicted Label: Against*

- Sometimes, the responder of a tweet disagrees with the news author, not with its publisher. Therefore, it is difficult to recognize the stance from the reply content. For example, if he uses insulting words in reply content, the algorithm assumes that he is in opposition, even though the person may be completely in agreement with the author of the tweet. See another sample in Table 10.

TABLE 10: ANOTHER SAMPLE OF DATASET

Main Post	Reply Post
After killing people in Goharshad Mosque, Reza Khan's agents took the bodies of the martyrs and even the wounded by truck to several places in Mashhad and buried them collectively. How do you really defend this executioner?	God damn him and his agents.
Real Label: Support	*Predicted Label: Against*

I. CONCLUSION AND FEATURE WORKS

Online platforms such as Twitter, Facebook, and discussion forums, have become popular platforms to discuss and express opinions about various topics. In this context, the stance is an opinion expressed by an individual towards some topic or issue, or personality. Stance detection aims to identify an author's stance towards a specific topic which has become a critical component in applications such as fake news detection, claim validation, author profiling, etc.

In this paper, we introduce a Persian stance detection dataset that can be used for several tasks such as account profiling, author personality detection, fake news detection, etc. Our dataset is composed of 3300 tweets (including main posts and their replies) collected from Twitter. Also, we applied a baseline model on this dataset which uses the ParsBert transformer. It should be noted that the announced results are primary and improvements can be made in the future.

In addition, we intend to use this stance classifier to build an end-to-end author personality detection. Another thing that can be done in the future is to build a larger dataset to train the model better. Also, due to having strong datasets in the English language, we can use cross-lingual methods for stance detection.

Another activity that can be done in the future is weighting stances. In other words, the stance weight of a person or a valid account that has a high influence on the social network should be higher than compared to a normal account.

Appendix: Dataset Labeling Guideline

The purpose of this guideline is to determine whether the author of an original post and a reply post agree on the main topic expressed in the original post. The following concepts are initially defined:

1. Original post: A post published by an author in the Twitter space and not in response to any other post.

[DOI: 10.52547/itrc.14.4.46]

2. Response post: A post by another author in response to the original post and probably contains another person's comments about the original post.
3. Post topic: The main topic that the author of a post had in mind to write a post, which may not be explicitly stated in the text, but it can be understood from the content of the post.
4. Author's stance on a topic: In addition to the topic of the post, the author of a tweet will probably have a stance on that topic, such as: support/agree, Deny/disagree and neither. In the following, we will describe each one:

- Support: The stance of the author of main post and the author of reply post to the subject of the post is the same regarding the topic of the original post.

Main Post	Reply Post	Stance
Ahwaz steel workers went on strike for a third day and gathered in front of the Khuzestan governor's office, protesting against low wages and job insecurity.	We need fundamental change, the lack of "job security" is the most important concern of workers.	Support

- Against: The stance of the author of main post and the author of reply post to the subject of the post is not the same.

Main Post	Reply Post	Stance
The US space agency, NASA, has released a report on the effects of global warming that some parts of the world, including Iran, will not be viable for another 30 years.	Raise your life expectancy. We are going to see happy days. For drought, it can be solved if the work is left to the skilled	Against

- Neither: It is not possible to detect whether the stance of the two people is the same or not.

Main Post	Reply Post	Stance
During the years 1393 to 1400, on average, about 4% of electricity consumption was saved due to changes in working hours of offices, which is equivalent to 106 billion kilowatt hours over 7 years.	Excuse me, do you have a scientific reference for this issue?	Neither

-Delete: This label should be selected in the following cases:

- tweets that contained insulting words.
- The tweet or its reply should not be in Farsi
- language The meaninglessness of the tweet or its response
- Insulting the members of the Islamic Republic of Iran

References

[1] Bhatt, Shaily, Naman Goenka, Sakshi Kalra, and Yashvardhan Sharma. "Fake News Detection: Experiments and Approaches beyond Linguistic Features." In Data Management, Analytics and Innovation, pp. 113-128. Springer, Singapore, 2022.

[2] Pomerleau, Dean, Delip Rao, the fake news challenge: Exploring how artificial intelligence technologies could be leveraged to combat fake news, 2017

[3] Mrowca, Damian, Elias Wang, and Atli Kosson. "Stance detection for fake news identification." Eliaswang. Com (2017).

[4] Mohammad, Saif, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. "Semeval-2016 task 6: Detecting stance in tweets." In Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), pp. 31-41. 2016.

[5] Schiller, Benjamin, Johannes Daxenberger, and Iryna Gurevych. "Stance detection benchmark: How robust is your stance detection?" KI-Künstliche Intelligenz 35, no. 3 (2021): 329-341.

[6] Mohammad, Saif, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. "A dataset for detecting stance in tweets." In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), pp. 3945-3952. 2016.

[7] Ferreira, William, and Andreas Vlachos. "Emergent: a novel data-set for stance classification." In Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: Human language technologies. ACL, 2016.

[8] Sobhani, Parinaz, Diana Inkpen, and Xiaodan Zhu. "A dataset for multi-target stance detection." In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers, pp. 551-557. 2017.

[9] Derczynski, Leon, Kalina Bontcheva, Maria Liakata, Rob Procter, Geraldine Wong Sak Hoi, and Arkaitz Zubiaga. "SemEval-2017 Task 8: RumourEval: Determining rumour veracity and support for rumours." arXiv preprint arXiv:1704.05972 (2017).

[10] Swami, Sahil, Ankush Khandelwal, Vinay Singh, Syed Sarfaraz Akhtar, and Manish Shrivastava. "An english-hindi code-mixed corpus: Stance annotation and baseline system." arXiv preprint arXiv:1805.11868 (2018).

[11] Gorrell, Genevieve, Elena Kochkina, Maria Liakata, Ahmet Aker, Arkaitz Zubiaga, Kalina Bontcheva, and Leon Derczynski. "SemEval-2019 task 7: RumourEval, determining rumour veracity and support for rumours." In Proceedings of the 13th International Workshop on Semantic Evaluation, pp. 845-854. 2019.

[12] Conforti, Costanza, Jakob Berndt, Mohammad Taher Pilehvar, Chryssi Giannitsarou, Flavio Toxvaerd, and Nigel Collier. "Will-They-Won't-They: A Very Large Dataset for Stance Detection on Twitter." arXiv preprint arXiv:2005.00388 (2020).

[13] Li, Yingjie, Tiberiu Sosea, Aditya Sawant, Ajith Jayaraman Nair, Diana Inkpen, and Cornelia Caragea. "P-stance: A large dataset for stance detection in political domain." In Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, pp. 2355-2365. 2021.

[14] Zarharan, Majid, Samane Ahangar, Fateme Sadat Rezvaninejad, Mahdi Lotfi Bidhendi, Mohammad Taher Pilevar, Behrouz Minaei, and Sauleh Eetemadi. "Persian Stance Classification Data Set." In TTO. 2019.

[15] Du, Jiachen, Ruifeng Xu, Yulan He, and Lin Gui. "Stance classification with target-specific neural attention networks." International Joint Conferences on Artificial Intelligence, 2017.

[16] Lai, Mirko, Alessandra Teresa Cignarella, Delia Irazú Hernández Farías, Cristina Bosco, Viviana Patti, and Paolo Rosso. "Multilingual stance detection in social media political debates." Computer Speech & Language 63 (2020): 101075.

[DOI: 10.52547/itrc.14.4.46]

[17] Zotova, Elena, Rodrigo Agerri, Manuel Nuñez, and German Rigau. "Multilingual stance detection in tweets: The Catalonia independence corpus." In Proceedings of the 12th Language Resources and Evaluation Conference, pp. 1368-1375. 2020.

[18] Darwish, Kareem, Walid Magdy, and Tahar Zanouda. "Trump vs. Hillary: What went viral during the 2016 US presidential election." In International conference on social informatics, pp. 143-161. Springer, Cham, 2017.

[19] Derczynski, Leon, Kalina Bontcheva, Michal Lukasik, Thierry Declerck, Arno Scharl, Georgi Georgiev, Petya Osenova et al. "Pheme: Computing veracity—the fourth challenge of big social data." In Proceedings of the Extended Semantic Web Conference EU Project Networking session (ESCW-PN). 2015.

[20] Bar-Haim, Roy, Indrajit Bhattacharya, Francesco Dinuzzo, Amrita Saha, and Noam Slonim. "Stance classification of context-dependent claims." In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers, pp. 251-261. 2017.

[21] Qazvinian, Vahed, Emily Rosengren, Dragomir Radev, and Qiaozhu Mei. "Rumor has it: Identifying misinformation in microblogs." In Proceedings of the 2011 conference on empirical methods in natural language processing, pp. 1589-1599. 2011.

[22] Hamidian, Sardar, and Mona T. Diab. "Rumor detection and classification for twitter data." arXiv preprint arXiv:1912.08926 (2019).

[23] Cignarella, Alessandra Teresa, Mirko Lai, Cristina Bosco, Viviana Patti, and Rosso Paolo. "Sardistance@ evalita2020: Overview of the task on stance detection in italian tweets." EVALITA 2020 Seventh Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (2020): 1-10.

[24] Lozhnikov, Nikita, Leon Derczynski, and Manuel Mazzara. "Stance prediction for russian: data and analysis." In International Conference in Software Engineering for Defence Applications, pp. 176-186. Springer, Cham, 2018.

[25] Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. "Attention is all you need." Advances in neural information processing systems 30, 2017

[26] Farahani, Mehrdad, Mohammad Gharachorloo, Marzieh Farahani, and Mohammad Manthouri. "Parsbert: Transformer-based model for persian language understanding." Neural Processing Letters 53, no. 6, 2021

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

[28] Kingma, Diederik P., and Jimmy Ba. "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980, 2014

[29] Abadi, Martin, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado et al. Tensorow: Largescale machine learning on heterogeneous distributed systems, arXiv preprint arXiv:1603.04467, 2016

[30] Küçük, Dilek, and Fazli Can. "A Tutorial on Stance Detection." In Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining, pp. 1626-1628. 2022.

[31] Wojatzki, Michael, and Torsten Zesch. "ltl. uni-due at semeval-2016 task 6: Stance detection in social media using stacked classifiers." In Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), pp. 428-433. 2016.

[32] Cignarella, Alessandra Teresa, Mirko Lai, Cristina Bosco, Viviana Patti, and Rosso Paolo. "Sardistance@ evalita2020: Overview of the task on stance detection in italian tweets." EVALITA 2020 Seventh Evaluation Campaign of Natural Language Processing and Speech Tools for Italian (2020): 1-10.

[33] Wei, Penghui, Junjie Lin, and Wenji Mao. "Multi-target stance detection via a dynamic memory-augmented network." In The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval, pp. 1229-1232. 2018.

[34] Zhou, Yiwei, Alexandra I. Cristea, and Lei Shi. "Connecting targets to tweets: Semantic attention-based model for target-specific stance detection." In International Conference on Web Information Systems Engineering, pp. 18-32. Springer, Cham, 2017.

[35] Tutek, Martin, Ivan Sekulić, Paula Gombar, Ivan Paljak, Filip Čulinović, Filip Boltužić, Mladen Karan, Domagoj Alagić, and Jan Šnajder. "Takelab at semeval-2016 task 6: Stance classification in tweets using a genetic algorithm based ensemble." In Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016), pp. 464-468. 2016.

[36] https://en.wikipedia.org/wiki/BERT_(language_model)

[37] Aharoni, Ehud, Anatoly Polnarov, Tamar Lavee, Daniel Hershcovich, Ran Levy, Ruty Rinott, Dan Gutfreund, and Noam Slonim. "A benchmark dataset for automatic detection of claims and evidence in the context of controversial topics." In Proceedings of the first workshop on argumentation mining, pp. 64-68. 2014.

[38] Rajendran, Gayathri, Prabaharan Poornachandran, and Bhadrachalam Chitturi. "Deep learning model on stance classification." In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 2407-2409. IEEE, 2017.

[39] Tshimula, Jean Marie, Belkacem Chikhaoui, and Shengrui Wang. "A pre-training approach for stance classification in online forums." In 2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), pp. 280-287. IEEE, 2020.

Mohammad Hadi Bokaei received the B.Sc. and M.Sc. degrees from Iranian University of Science and Technology (2008) and Sharif University of Technology (2011), respectively, and the Ph.D. degree in Artificial Intelligence from Sharif University of Technology in 2015. He is currently an Associate Professor at the Iran Telecommunication Research Center, Tehran, Iran. His research interests include the area of Deep Learning, Machine Learning, Spoken Language Processing and Natural Language Processing.

Mojgan Farhoodi received her B.Sc. degree in Software Engineering and M.Sc. degree in IT Engineering from Amirkabir University of Technology. She has been working as a researcher at the ICT Research Institute (ITRC) since 2010. Currently, he is working as a faculty member of the research institute in the fields of Information Retrieval, Natural Language Processing, and Artificial Intelligence.

[DOI: 10.52547/itrc.14.4.46]

Mona Davoudi Shamsi received her B.Sc. degree from Shahid Beheshti University in the field of Software Engineering. Since 1380, he has been engaged in research activities at the ICT Research Institute (ITRC). In recent years, he has been working as a researcher in fields related to Natural Language Processing.

[DOI: 10.52547/itrc.14.4.46]