Claim Detection in Persian Twitter Posts

|  |  |  |
| --- | --- | --- |
| Mohammad Hadi Bokaei ICT Research InstituteTehran, Iranmh.bokaei@itrc.ac.ir | Minoo Nassajianminoonassajian@alum.sharif.edu | Mona Davoudi ShamsiICT Research InstituteTehran, Irandavoudi@itrc.ac.ir |

***Abstract—The considerable growth of false information on social media has harmful consequences on all aspects of people's lives. There have been conducted various studies on designing automated fact-checking systems to promote the veracity and correctness of news and information to reduce the harmful effects of fake news and misinformation. Claim detection is regarded as the first step of developing fact-checking systems on which some studies have been done in a few languages. In the current paper, we aim to provide a corpus of Persian tweets and analyze them linguistically for annotation. Moreover, we develop a baseline claim detection model to evaluate the dataset. This study is framed as a classification task and enjoys transformer-based method to train a multi-label classifier to detect the different types of claims in Persian texts.***

***Keywords- Automatic Claim Detection, Persian Language Processing, Low-resource language, Multi Label Classification***

# Introduction

Social media is considered as the most important platform for quickly and easily disseminating information nowadays. It paves the way for people to spread information in various formats such as texts, pictures, videos, and audios in the shortest time. Users can access to different news sources with a cellphone and simply share them with others.

In addition to a considerable sum of benefits, human achievements in technology have also brought disadvantages. The web contains a huge amount of information, some of which is incorrect and some sources present only one side of a contentious issue. This problem is regarded as a growing issue in public, [3], [4] as incorrect information can manipulate people's perceptions of reality, conscious and unconscious attitudes, and behaviors [5], and causes a mistrust of various problems [6] (especially during crisis [7]) among societies from different aspects such as health behavior [8], [9], [10], [11], political attitudes and voting behaviors [12], [13], financial markets [14], [15], [16], cognitive psychology [17], [18], [19] etc. Journalists and fact-checkers constantly work to check the veracity of released information and news and correct misinformation. As AI-based tools can alleviate the burdensome and time-consuming human activities, there have been conducted various studies on presenting fact-checking tools [20], [21], [22]. Fact-checking provides an accurate and unbiased analysis of claims in order to improve individuals' knowledge of important issues [23], [24]. A claim is a statement or assertion about something typically without providing evidence or proof, and it is widely considered the conclusive and central part of an argument [25], [26]. Claim detection is considered as the first step in the automated fact-checking pipeline [27]. Although more studies have been conducted on fact-checking with different names in literature like disinformation detection, fake news detection, or rumor detection, claim detection is a new study in natural language processing (NLP) that has attracted the attention of researchers in recent years.

In this paper, we set out to present the first dataset of claims from Persian tweets. We used an annotation schema that effectively captures the different types of claims and non-claim based on linguistic analyses. The dataset contains 4910 tweets labeled by 2 persons. This corpus has been used in the development of an automated claim 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 the different types of claims to classify Persian tweets. Section 4 introduces our methodology and corpus information. In section 5, the experimental reports are presented, including evaluation metrics, error analysis, and results. Section 6 concludes the survey and suggests future directions in this area.

# Literature Review

The claim detection is regarded as a sub-task in the wider field of argumentation mining (AM) [30] and fact checking task [32]. In these tasks, analyzing claims is regarded as the main part to provide structured data for computational models of arguments and reasoning engines [31] (in AM) and to assessing the truthfulness of claims (in fact-checking) [32]. [33], [34], [35] have discussed automated fact-checking process in the context of computational journalism. This study has recently received significant attention in the field of NLP as well. In AI-based research, the automated fact-checking process consists of three stages the first of which is claim detection [36]. Some methods of automated fact checking assume that a claim is given as input, however in the context of fully automated fact checking this will not be the case, as it is designed to work on articles or on comment sections of social media. So, the first step must be to determine what is and is not a claim [37].

A few studies have been recently done on claim detection in a limited number of languages such as English [48], [49], [50], [51], Arabic [52], Turkish [53], [55], and Dutch [56]. Most of these studies have been framed as a classification task. For example, [48] used a multi-class SVM classifier, [57] used a deep model with a Feed-Forward Neural Network (FNN) as the last layer to rank labels, and [58] and [59] used transformer-based models. However, there are some works used sequence labeling techniques [60], [61] and an unsupervised learning approach [62]. In the mentioned works, two approaches were applied to annotated datasets. According to the first approach, datasets were provided relying on the concept of check-worthiness and/or importance of claims, and three labels were defined to classify sentences into one of 1) claim, 2) non-claim, and 3) unimportant claim. In the second approach, researchers avoid the subjective concepts of check-worthiness or importance and they left it to fact-checker. They linguistically analyzed sentences and defined multi labels to determine different types of claims [26].

In Persian, as one of the low-resource languages, there is not any research on automatic claim detection. There are just a few studies on stance detection [64], fake news detection [65] and rumor verification [69] as related studies. So, the present paper focuses on the Persian claim detection task and developing a dataset for it for the first time.

# Dataset

To provide an annotation guidance for labeling our corpus, we use 4910 Persian tweets and analyze it linguistically. Against some studies relying on the concept of check-worthiness of claims resulting in subjective interpretation, we ignore the judgment of importance and left it to factcheckers and journalists, and consider 10 labels to categorize tweets as different types of claim and one label for the type of non-claim. In comparison with previous studies such as [27] that defined 19 sub-categories for the types of claims, we just define 10 labels for the types of claims with high frequencies and classify the types of claims with low frequencies as the other claim label. As a tweet can consist of several sentences or the different types of information, it can be annotated by multi labels. In the next part, we introduce the sub-categorization of claims.

## The sub-categorization of claims

To categorize claims, we use linguistic features and define 10 classes (Action, Prediction, Support/Oppose, Causation/Correlation, Quantity, Comparison, Quote, Trait, Law/Rule, Other Claim) based on syntactic, semantic, and pragmatic analysis. A non-claim label is also defined for a statement that is not a claim. We will describe each categorization as follow.

### Action claims

This type of claim encompasses actions performed in the past or are currently being taken. In addition, events happening in the near future are considered as this type. The below examples respectively show the past (as shown in example 1 and 4), present (as in example 2), and future tenses (as seen in example 3) in Persian that are annotated by the action claim label.

1. *500 pezes̄k sāle gozas̄te māliyāt pardāxt nakardand.*

“500 doctors did not pay tax last year”

1. *In kes̄varhā dar hāle tote'ec̄ini barāye hamle be suriye hastand.*

“These countries are planning to attack Syria.”

1. *Fardā tahrimhā e`māl xāhad s̄od.*

“Sanctions will be imposed tomorrow.”

1. *Agar budje be s̄erkat taxsis miyāft, sāzmān vars̄ekast nemis̄od.*

“If the budget had been allocated to the company, it would not have gone bankrupt.”

In addition to the above structures, claims can be made in the forms of presuppositional structures. According to 6 different types of presupposition [72], active presupposition and lexical presupposition consist of words (such as *‘know’*, *‘regret’*, *‘realize’*, *‘start’*, *‘stop’*, *‘again’* etc.) showing an assumption about taking an action in the past. Examples 5 and 6 presuppose actions taken only in the past.

1. *In kes̄var mojaddad e`māle tahrimhā rā elayhe Irān āgāz kard.*

“This country again imposing sanctions against Iran starts”

1. *Dolat digar be panāhandegān ejāzeye vorud nemidahad.*

“The government no longer allows refugees to enter.”

### Prediction claims

Prediction claims include syntactic and semantic structures predicting events in the future. These kinds of simple and complex sentences (exemplified in 7 to 9) can contain adverbs for future (such as *by the end of this week/month/year, soon, in 2 days, etc*.). Moreover, some expressions or words showing a prediction or an expectation (such as *it is predicted/ expected that..., it is possible/likely...., etc.*) given in 10. Future tense (as seen in 7), present tense (illustrated in 8 and 9), and subjunctive forms (as shown in 10 and 11) are used for Persian verbs to show taking actions in the future.

1. *Mo`āmelāte bāzāre sahām tā pāyāne sāle jāri tahte ta`sire tavarrom qarār xāhad gereft.*

“Stock market transactions will be affected by inflation by the end of this year.”

1. *Bā taxsise budje ta 2 māhe āyande proje rā be etmām miresānim.*

“By allocating the budget, we will complete the project in the next 2 months.”

1. *Agar budje taxsis yābad, tā 2 māhe āyande proje rā be etmām miresānim.*

*“*If the budget is allocated, we will complete the project in the next two months.”

1. *Pi`s̄bini mis̄avad pis̄rafte in proje emsāl be bis̄ az 60% beresad.*

“The progress of this project is expected to reach more than 60% this year.”

1. *Ehtemāl dārad āmrikā dar āyandeye nazdik bā jange dāxeli movājeh s̄avad.*

“It is possible America face a civil war in the near future.”

### Support/Oppose claims

This claim includes statements supporting, opposing, or remaining neutral on an issue or a person's opinion (in 12 to 14).

1. *Rand Paul mox**ālefe tavāfoqe haste`i bā Iran ast.*

“Rand Paul opposes the nuclear deal with Iran.”

1. *Demokrāthā be lāyeheye zirsāxtha ra`ye moxālef dādand.*

“Democrats voted against the infrastructure bill.”

1. *Donald Trump farmāne ejrā`i e`māle tahrimhāye jadid alayhe Irān rā emzā kard.*

“Donald Trump signed an executive order imposing new sanctions on Iran.”

### Causation/Correlation claims

This type of claim aims to capture sentences asserting at least 2 events occurring. In causation claims, one event causes occurring another one (as shown in 15 to 17). In correlation claims, there is a correlation between two events (as given in 18). To make this claim, if-then structures, prepositional phrases to express one of the events, and causative verbs can be used in Persian.

1. *Agar budje be s̄erkat taxsis miyāft, sāzmān varsekast nemisod.*

“If the budget had been allocated to the company last year, it would not have gone bankrupt.”

1. *40% az s̄erkathā bā e`māle tahrimhā vars̄ekast sodand.*

“40% of companies have gone bankrupt due to sanctions.”

1. *Vāksane koronā bā`ese nābārvari mis̄avad.*

“Covid-19 vaccine causes infertility.”

1. *Har zamān haddeaqal hoquq rā afzāyes̄ dādim, ros̄di dar mas̄āqel mos̄āhede s̄od.*

“Every time we've increased the minimum wage, we've seen a growth in jobs.”

### Quantity claims

This subcategory encompasses ratio and percentage, ranking, date, and numerical and statistical analyses. The below examples (19 to 22) show this kind of claim.

1. *Nerxe bikāri dar Irān sale gozas̄te 9.5 darsad bud.*

“Last year, unemployment rate was 9.5 percent in Iran.”

1. *Mā dovvomin sāderkonandeye naft dar jahān hastim.*

“We are the second oil exporter in the world.”

1. *Ānhā 10 sāl māliyāt pardāxt nakardeand.*

“They have not paid taxes for 10 years.”

1. *Tahrimhā 40 miliyārd dolār manābe`e arzi rā masdud kard.*

“Sanctions froze $ 40 billion worth of foreign currencies.”

### Comparison claims

This subcategory includes all comparative structures such as the comparison between 2 or more things (as seen in 23), relationship between qualities in/over time (in 24), uniqueness (as given in 25), similarity and difference (exemplified in 26).

1. *Espāniyā bis̄tarin tedāde javānāne bikār rā dārad.*

“Spain has the most unemployed young people.”

1. *Nerxe bikāri dar dolate fe`li kamtar az dolate qabli ast.*

“The unemployment rate in the current government is lower than that in the previous government.”

1. *Āmrikā tanhā kesvari ast ke nerxe bikāri dar ān sefr ast.*

“America is the only country where the unemployment rate is zero."

1. *Bar xalāfe dolate qabli Nerxe bikari dar dolate fe`li xeili pāyin ast.*

“Unlike the previous government, the unemployment rate in the current government is very low.”

### Quote claims

This class encompass claims that repeat or paraphrase what an entity said. The following examples show a quote claim and its paraphrase.

1. *Trump goft: tahrimhāye jadid bānke melli Irān rā emruz e`māl mikonim.”*

“Trump said: we impose new sanctions against Iran's national bank today.”

1. *Ra`is jomhure āmrikā e`lām kard tahrimhāye jadidi alayhe bānke melli Irān emruz e`māl mis̄avad.*

Figure 1: The distribution of 11 classes labeled by annotator 1 and annotator

“The president of America said new sanctions against Iran's national bank will be imposed today.”

### Trait claim

This type of claim covers different properties of an entity such as strength, weakness, capability, qualification etc. using especial verbs such as “able, be capable, can (as seen in 29) and modifiers such as adjective phrases (as in 30).

1. *In kes̄var qāder ast dar āyandeh`i nazdik be selāhe haste`i dast yabad.*

“This country is capable of acquiring nuclear weapons in the near

future.”

1. *Is̄ān za'iftarin ra'is jomhure tārixe in kesvar hastand.*

“He is the weakest president in the history of this country."

### Law/Rule claim

This type of claim contains statements that express laws and regulations and consider actions permissible and impermissible. Examples 31 and 32 show this kind of claim.

1. *Dolat be s̄erkathāye xāreji ejāze midahad dar bāzar mos̄ārekat konand.*

“The government allows foreign companies to participate in its market.”

1. *Sāxtosāz dar in rustā qeire qānuni ast.*

“The constructions of this village are illegal.”

### Other claims

This type of claim contains statements that do not fit into any of the previous categories (as given in 33 and

1. *Āmāre bikārān c̄āles̄barangiz va negarānkonande ast.*

“Unemployment statistics are challenging and worrying.”

1. *Hadafe mā jang nist.*

“Our goal is not war.”

### Non-claim

There are some tweets such as people's personal opinion, personal experiences, advises, poems etc. that are not considered as claim. We define a non-claim label to annotate these sentences (as exemplified in 35 and 36).

1. *Sāle no mobārak.*

“Happy new year.”

1. *Inbār behtarin tajrobeye safaram ra dās̄tam.*

“I had the best travel experience this time.”

# Methodology

To provide the diversity of claims, at first, we used 120 tweet accounts. We filtered these tweets based on their contents. We removed tweets containing non-Persian texts. Moreover, we considered a threshold for the length of tweets and removed sentences shorter than 8 words. Furthermore, we collect tweets that have a considerable number of retweet (more than 30 retweets) and like (more than 30 likes). Finally, we used the 60 most followed public figures’ accounts on tweeter, including Iranian politicians, news and television presenters, actors, and famous social media influencers. We collected their tweets posted between May 2016 and January 2020. The shortest and longest tweets in the corpus consist of 8 and 70 tokens respectively and the average tweet length is around 40 tokens. In this paper, we applied the classification technique to model claim detection task in both binary and multi-label classification. The following parts will present detailed information on corpus statistics and the model.

## Data and Annotation

To annotate tweets, we used doccano, an open-source data labeling tool for machine learning practitioners. This platform supports collaborative annotations and different languages. We provided a preliminary set of annotation guideline based on linguistically analyzing 200 tweets and went through three iterations of refinements to reach to the final annotation guideline. Two annotators were given the guideline, detailed explanations, and examples of 11 categories of claims. The agreement percentage among labels is 33.90%, and the agreement percentage with the minimum of one agreement is 64.69%. The Cohen's Kappa static [73] is 0.58 for the binary claim/non-claim annotation task. The Krippendorff’s alpha [74] was also calculated for the binary case and the results is 0.57. These results are better and more reasonable than [27]. In fact, their method for mapping seven claim categories into the binary category is just based on reaching to a higher inter-annotator agreement. They mapped some claim types into not a claim group in binary case to get a higher value of alpha, whereas we mapped all types of claims into a claim class in the binary classification. Figure 1 presents the distribution of 11 annotated classes by annotator 1 and 2 and figure 2 shows the distribution of tweets in a binary class. According to the figures, we can see that both annotators have a high agreement on distinguishing the claim class from the not a claim class, and the most of the disagreement was on the causation/correlation class.

# Experiments

To evaluate the dataset, we train a model to distinguish claims from nonclaims in both binary and multi-label classifications. In this section, we first talk about the model and experimental setting. Then, the results of each classifier in both binary and multi-label settings are reported. Finally, the errors of the model are discussed in order to open the way for the future studies of the topic.

##  Model

 We experiment with a transformer-based [75] model for each of binary and multi-label classification tasks, called ParsBert [76]. This model is a monolingual language model for Persian language with the same configurations as Bert [77], pre-trained on different texts such as news, novels, scientific documents etc. We fine-tuned this model using the claim tweet corpus. It is followed by a fully-connected network to map the ParsBert's outputs to the tag space.

## Experimental Setting

The number of tokens and tweets in the corpus are 197480 and 4910 respectively. 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 [78] was applied for optimizing the model. We used TensorFlow library [79] 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 11 sigmoid functions as the output layer to predict the labels of a tweet. Moreover, we used binary cross entropy as the loss function.

Figure 2: Distribution of binary class (claim/non-claim)

##  Results

Table 1 reports the results of both binary and multi-label classifications on the test set. In the binary classification task, we consider 2 gold tag sets. The first one contains the union of annotators' labels and the second one includes the intersection of annotators' labels. In addition to these gold tag sets, we compare the predicted labels with each annotator's tag set in the multilabel classification task. According to the results, it can be seen that the intersection label task enjoyed a more balanced precision and recall and obtained the highest F-score in the binary setting, 89.34%. In the multilabel classification, the intersection label task also outperformed the union label task by 0.45%.

## Analysis

From the best performance of the model in both binary and multi-label settings, we conclude that the model could recognize sentences with the content of advice and correctly labeled them non-claim, as given in (a).

*(a) Din ādam rā mostaqel bār miāvarad va u rā ros̄d midahad.*

“Religion makes a person independent and develops him.”



Table 1: The results for claim detection experiments, separated into binary and multi-label evaluations. The best F1 score is printed in bold face.

As transformer-based models enjoy contextualized embeddings that capture syntactic relations [80, 81, 82], we also see that the model can recognize different syntactic structures well. The model can capture causative structures and future tenses and classify them as causation and prediction claims correctly. As causative structures express an action which is caused to happen, the model can consider the action claim label for causative sentences as well. For example, the model correctly labeled (b) as causative/correlation and action claims.

(*b)Vaz'iyate bohrāni dar in s̄ahr natijeye adame tavajoh be hos̄dārhā ast!*

“The critical situation in this city is the result of not paying attention to the warnings!”

In addition, the model uses morphological features well to capture comparative and superlative adjectives to correctly categorize sentences as comparison claims (c).

*(c)Tasmimāte ra'is jomhure fe'li bohrāne eqtesādi rā badtar mikonad.*

“The current president's decisions are worsening economic crisis.”

Moreover, words and punctuation showing quotations or paraphrases could be correctly captured by the model to recognize quote claims, as seen in (d) and (e).

*(d) Irnā: 30% as bimārāne viruse koronā dar in s̄ahr mosāfer hastand.*

“Irna: 30% of COVID-19 patients in this city are travelers.”

*(e) Keyhān neves̄t ke xabarnegāre panāhande eslāhāt talab ast.*

“Kayhan wrote that the refugee journalist is a reformist.”

Law/Rule claim could be detected by the model well. The model could recognize words expressing rules in tweets (f).

*(f) Tarhe dolat barāye sāderāte xodro tasvib sod.*

“The government's plan to export cars was approved.”

The model could well capture words that express supporting and opposing, as seen in (g).

*(g) Namāyandegān az tarhe jadide dolat hemāyat kardand.*

“Members of parliaments supported the government's new plan.”

Tense as a syntactic feature could be captured really well and this feature helped the model to detect prediction claims, as given in (h).

*(h) In kes̄var az haqe mardome xod darbāreye mozākerāte haste'i kutāh naxāhad āmad.*

“This country will never step back from the rights of their people in nuclear negotiations."

In spite of the aforementioned points, there are some frequent and important errors in the performance of the model. Although the attention heads in transformer-based models could well capture the syntactic structures of sentences [80, 81], the model made several errors:

1. Considering morphological features, the model labeled some tweets containing ordinal numbers as quantity claim. For instance, (i) is not a claim but the model's tag is quantity claim.

*(i) Emruz noxostin ruze qarne pānzdahom ast.*

“Today is the first day of the 15th century."

2. As we mentioned in section 3.1.4, causative structures can be formed by if-then structures in Persian. However, all if-then structures are not causative. The model incorrectly labeled (j) as causation/correlation claim.

*(j) Agar tavāne moqābele ba vaz'iyate bohraniye fe'li ra nadārid, este'fā dahid.*

“If you cannot cope with the current critical situation, resign."

3. There are some tweets incorrectly labeled by annotators. However, the model could correctly label them. For instance, the model correctly tagged (k) as trait claim.

*(k) Modire bānk ideye eqtesādi nadārad.*

“The bank manager has no economic idea.”

# Conclusion and Future Work

We have developed the first annotation schema for Persian claim detection based on linguistic analyses. We created an annotated dataset made of sentences extracted from Persian tweets posted by Iranian public figures. We experimented with a transformer-based model, ParsBert, in the tasks of binary and multi-label claim detection. The results show that the model could well captures syntactic features to detect the types of claims. However, the model has some weak points to semantically analyze sentences in order to identify the types of claim in sentences with the same syntactic structures. This study can be used in developing the next steps of fact-checking pipeline like stance detection.

##### References

[1] Herman, Edward S., and Noam Chomsky. Manufacturing consent: The political economy of the mass media. Random House, 2010.

[2] Bradshaw, Samantha, and Philip N. Howard. "The global disinformation order: 2019 global inventory of organised social media manipulation." (2019).

[3] Shu, Kai, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. "Fake news detection on social media: A data mining perspective." ACM SIGKDD explorations newsletter 19, no. 1 (2017): 22-36.

[4] Zubiaga, Arkaitz, Ahmet Aker, Kalina Bontcheva, Maria Liakata, and Rob Procter. "Detection and resolution of rumours in social media: A survey." ACM Computing Surveys (CSUR) 51, no. 2 (2018): 1-36.

[5] Bastick, Zach. "Would you notice if fake news changed your behavior? An experiment on the unconscious effects of disinformation." Computers in human behavior 116 (2021): 106633.

[6] Ecker, Ullrich KH, Stephan Lewandowsky, John Cook, Philipp Schmid, Lisa K. Fazio, Nadia Brashier, Panayiota Kendeou, Emily K. Vraga, and Michelle A. Amazeen. "The psychological drivers of misinformation belief and its resistance to correction." Nature Reviews Psychology 1, no. 1 (2022): 13-29.

[7] Plotnick, Linda, Starr Hiltz, Sukeshini Grandhi, and Julie Dugdale. "Real or fake? User behavior and attitudes related to determining the veracity of social media posts." arXiv preprint arXiv:1904.03989 (2019).

[8] Featherstone, Jieyu Ding, C. A. Davis, and J. Zhang. "Correcting Vaccine Misinformation on Social Media Using Fact-checking Labels."

In APHA's 2019 annual meeting and expo. 2019.

[9] Roozenbeek, Jon, Claudia R. Schneider, Sarah Dryhurst, John Kerr, Alexandra LJ Freeman, Gabriel Recchia, Anne Marthe Van Der Bles, and Sander Van Der Linden. "Susceptibility to misinformation about COVID-19 around the world." Royal Society open science 7, no. 10 (2020): 201199.

[10] Greene, Ciara M., and Gillian Murphy. "Quantifying the effects of fake news on behavior: Evidence from a study of COVID-19 misinformation." Journal of Experimental Psychology: Applied (2021).

[11] Zhou, Cheng, Haoxin Xiu, Yuqiu Wang, and Xinyao Yu. "Characterizing the dissemination of misinformation on social media in health emergencies: An empirical study based on COVID-19." Information Processing and Management 58, no. 4 (2021): 102554.

[12] Van Der Linden, Sander, Costas Panagopoulos, and Jon Roozenbeek. "You are fake news: political bias in perceptions of fake news." Media, Culture and Society 42, no. 3 (2020): 460-470.

[13] Muqsith, Munadhil Abdul, R. Ridho Pratomo, Anna Gustina Zaina, and Ana Kuswanti. "Fake News as a Tool to Manipulate the Public with False Information." In 2nd International Indonesia Conference on Interdisciplinary Studies (IICIS 2021), pp. 118-127. Atlantis Press, 2021.

[14] Polak, Mateusz. "The misinformation effect in financial markets: An emerging issue in behavioural fianance." e-Finanse: Financial Internet Quarterly 8, no. 3 (2012): 55-61.

[15] Kogan, Shimon, Tobias J. Moskowitz, and Marina Niessner. "Fake news: Evidence from financial markets." Available at SSRN 3237763 (2019).

[16] Rich, Patrick R., and Maria S. Zaragoza. "The continued inuence of implied and explicitly stated misinformation in news reports." Journal of experimental psychology: learning, memory, and cognition 42, no. 1 (2016): 62.

[17] Thorson, Emily. "Belief echoes: The persistent effects of corrected misinformation." Political Communication 33, no. 3 (2016): 460-480.

[18] Desai, Saoirse Connor, and Stian Reimers. "Some misinformation is more easily countered: An experiment on the continued inuence effect." In CogSci. 2018.

[19] Wang, Xuezhi, Cong Yu, Simon Baumgartner, and Flip Korn. "Relevant document discovery for fact-checking articles." In Companion Proceedings of the The Web Conference 2018, pp. 525-533. 2018.

[20] Jo, Saehan, Immanuel Trummer, Weicheng Yu, Xuezhi Wang, Cong Yu, Daniel Liu, and Niyati Mehta. "Verifying text summaries of relational data sets." In Proceedings of the 2019 International Conference on Management of Data, pp. 299-316. 2019.

[21] Trokhymovych, Mykola, and Diego Saez-Trumper. "Wikicheck: An end-to-end open source automatic fact-checking api based on wikipedia." In Proceedings of the 30th ACM International Conference on Information and Knowledge Management, pp. 4155-4164. 2021.

[22] Graves, Lucas, Brendan Nyhan, and Jason Reier. "Understanding innovations in journalistic practice: A field experiment examining motivations for fact-checking." Journal of Communication 66, no. 1 (2016): 102-138.

[23] Graves, D. "Understanding the promise and limits of automated factchecking." (2018).

[24] Lippi, Marco, and Paolo Torroni. "Context-independent claim detection for argument mining." In Twenty-Fourth International Joint Conference on Artificial Intelligence. 2015.

[25] Stab, Christian, and Iryna Gurevych. "Parsing argumentation structures in persuasive essays." Computational Linguistics 43, no. 3 (2017): 619-659.

[26] Konstantinovskiy, Lev, Oliver Price, Mevan Babakar, and Arkaitz Zubiaga. "Toward automated factchecking: Developing an annotation schema and benchmark for consistent automated claim detection." Digital Threats: Research and Practice 2, no. 2 (2021): 1-16.

[27] Shin, Jieun, Lian Jian, Kevin Driscoll, and Francois Bar. "The diffusion of misinformation on social media: Temporal pattern, message, and source." Computers in Human Behavior 83 (2018): 278-287.

[28] Wu, Liang, Fred Morstatter, Kathleen M. Carley, and Huan Liu. "Misinformation in social media: definition, manipulation, and detection." ACM SIGKDD Explorations Newsletter 21, no. 2 (2019): 80-90.

[29] Levy, Ran, Yonatan Bilu, Daniel Hershcovich, Ehud Aharoni, and Noam Slonim. "Context dependent claim detection." In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics: Technical Papers, pp. 1489-1500. 2014.

[30] Lippi, Marco, and Paolo Torroni. "Argument mining: A machine learning perspective." In International Workshop on Theory and Applications of Formal Argumentation, pp. 163-176. Springer, Cham, 2015.

[31] Vlachos, Andreas, and Sebastian Riedel. "Fact checking: Task definition and dataset construction." In Proceedings of the ACL 2014 workshop on language technologies and computational social science, pp. 18-22. 2014.

[32] Daniel, Anna, Terry Flew, and Christina Spurgeon. "The promise of computational journalism." In Proceedings of the Australian and New Zealand Communication Association (ANZCA) Conference 2010: Media, Democracy and Change, pp. 1-19. Australia and New Zealand Communication Association, 2010.

[33] Cohen, Sarah, Chengkai Li, and Jun Yang. "C. Yu. Computational journalism: A call to arms to database researchers." CIDR, 2011.

[34] Graves, D. "Understanding the promise and limits of automated factchecking." (2018).

[35] Guo, Zhijiang, Michael Schlichtkrull, and Andreas Vlachos. "A survey on automated fact-checking." Transactions of the Association for Computational Linguistics 10 (2022): 178-206.

[36] Lazarski, Eric, Mahmood Al-Khassaweneh, and Cynthia Howard. "Using NLP for Fact Checking: A Survey." Designs 5, no. 3 (2021)

[37] Walton, Douglas. "Argumentation theory: A very short introduction." In Argumentation in artificial intelligence, pp. 1-22. Springer, Boston, MA, 2009.

[38] Moens, Marie-Francine, Erik Boiy, Raquel Mochales Palau, and Chris Reed. "Automatic detection of arguments in legal texts." In Proceedings of the 11th international conference on Artificial intelligence and law, pp. 225-230. 2007.

[39] Palau, Raquel Mochales, and Marie-Francine Moens. "Argumentation mining: the detection, classification and structure of arguments in text." In Proceedings of the 12th international conference on arti-ficial intelligence and law, pp. 98-107. 2009.

[40] Saint-Dizier, Patrick. "Processing natural language arguments with the< TextCoop> platform." Argument and Computation 3, no. 1 (2012): 49-82.

[41] Cabrio, Elena, and Serena Villata. "Natural language arguments: A combined approach." In ECAI 2012, pp. 205-210. IOS Press, 2012.

[42] Stylianou, Nikolaos, and Ioannis Vlahavas. "Transformed: End-to-End transformers for evidence-based medicine and argument mining in medical literature." Journal of Biomedical Informatics 117 (2021):103767.

[43] Al Khatib, Khalid, Tirthankar Ghosal, Yufang Hou, Anita deWaard, and Dayne Freitag. "Argument mining for scholarly document processing: Taking stock and looking ahead." In Proceedings of the Second Workshop on Scholarly Document Processing, pp. 56-65. 2021.

[44] Mebane, Waleed. "Detection of Claims and Supporting Evidence in Wikipedia Articles on Controversial Topics." PhD diss., 2017.

[45] Lawrence, John, and Chris Reed. "Combining argument mining techniques." In Proceedings of the 2nd Workshop on Argumentation Mining, pp. 127-136. 2015.17

[46] Budzynska, Katarzyna, Mathilde Janier, Juyeon Kang, Chris Reed, Patrick Saint-Dizier, Manfred Stede, and Olena Yaskorska. "Towards argument mining from dialogue." In Computational Models of Argument, pp. 185-196. IOS Press, 2014.

[47] Patwari, Ayush, Dan Goldwasser, and Saurabh Bagchi. "Tathya: A multi-classifier system for detecting check-worthy statements in political debates." In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pp. 2259-2262. 2017.

[48] Hassan, Naeemul, Gensheng Zhang, Fatma Arslan, Josue Caraballo, Damian Jimenez, Siddhant Gawsane, Shohedul Hasan et al. "Claimbuster: The first-ever end-to-end fact-checking system." Proceedings of the VLDB Endowment 10, no. 12 (2017): 1945-1948.

[49] Gencheva, Pepa, Preslav Nakov, Liuis Marquez, Alberto Barron Cede, and Ivan Koychev. "A context-aware approach for detecting worth-checking claims in political debates." In Proceedings of the International Conference Recent Advances in Natural Language Processing,

RANLP 2017, pp. 267-276. 2017.

[50] Hanto, Vigdis, and Mats Tostrup. "Towards Automated Fake News Classification-On Building Collections for Claim Analysis Research." Master's thesis, NTNU, 2018.

[51] Atanasova, Pepa, Preslav Nakov, Georgi Karadzhov, Mitra Mohtarami, and Giovanni Da San Martino. "Overview of the CLEF-2019 CheckThat! Lab: Automatic Identification and Verification of Claims. Task 1: Check-Worthiness." CLEF (Working Notes) 2380 (2019).

[52] Shaar, Shaden, Alex Nikolov, Nikolay Babulkov, Firoj Alam, Alberto Barron-Cedeno, Tamer Elsayed, Maram Hasanain et al. "Overview of CheckThat! 2020 English: Automatic identification and verification of claims in social media." In CLEF (Working Notes). 2020.

[53] Atanasova, Pepa, Alberto Barron-Cedeno, Tamer Elsayed, Reem Suwaileh, Wajdi Zaghouani, Spas Kyuchukov, Giovanni Da San Martino, and Preslav Nakov. "Overview of the CLEF-2018 CheckThat! Lab on automatic identification and verification of political claims. Task 1: Check-worthiness." arXiv preprint arXiv:1808.05542 (2018).

[54] Barron-codeno, Alberto, Tamer Elsayed, Preslav Nakov, Giovanni Da San Martino, Maram Hasanain, Reem Suwaileh, Fatima Haouari et al. "Overview of CheckThat! 2020: Automatic identification and verification of claims in social media." In International Conference of the Cross-Language Evaluation Forum for European Languages, pp. 215-236. Springer, Cham, 2020.

[55] Kartal, Yavuz Selim, and Mucahid Kutlu. "TrClaim-19: The first collection for Turkish check-worthy claim detection with annotator rationales." In Proceedings of the 24th Conference on Computational Natural Language Learning, pp. 386-395. 2020.

[56] Berendt, Bettina, Peter Burger, Rafael Hautekiet, Jan Jagers, Alexander Pleijter, and Peter Van Aelst. "FactRank: Developing automated claim detection for Dutch-language fact-checkers." Online Social Networks and Media 22 (2021): 100113.

[57] Atanasova, Pepa, Preslav Nakov, Luis Marquez, Alberto Barron-Cede~no, Georgi Karadzhov, Tsvetomila Mihaylova, Mitra Mohtarami, and James Glass. "Automatic fact-checking using context and discourse information." Journal of Data and Information Quality (JDIQ) 11, no. 3 (2019): 1-27.18

[58] Williams, Evan, Paul Rodrigues, and Valerie Novak. "Accenture at CheckThat! 2020: If you say so: post-hoc fact-checking of claims using transformer-based models." arXiv preprint arXiv:2009.02431 (2020).

[59] Hasanain, Maram, and Tamer Elsayed. "bigIR at CheckThat! 2020: Multilingual BERT for Ranking Arabic Tweets by Check-worthiness." In CLEF (Working Notes). (2020).

[60] Goudas, Theodosis, Christos Louizos, Georgios Petasis, and Vangelis Karkaletsis. "Argument extraction from news, blogs, and social media." In Hellenic Conference on Artificail Intelligence, pp. 287-299. Springer, Cham, (2014).

[61] Sardianos, Christos, Ioannis Manousos Katakis, Georgios Petasis, and Vangelis Karkaletsis. "Argument extraction from news." In Proceedings of the 2nd Workshop on Argumentation Mining, pp. 56-66. (2015).

[62] Levy, Ran, Shai Gretz, Benjamin Sznajder, Shay Hummel, Ranit Aharonov, and Noam Slonim. "Unsupervised corpus {wide claim detection." In Proceedings of the 4th Workshop on Argument Mining,

pp. 79-84. (2017).

[63] Josue, Caraballo. \PolitiTax A Taxonomy of Political Claims.”. (2018). https://https://perma.cc/4RQF-FCPV.

[64] 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.

[65] Samadi, Mohammadreza, Maryam Mousavian, and Saeedeh Momtazi. "Persian fake news detection: Neural representation and classification at word and text levels." Transactions on Asian and Low-Resource Language Information Processing 21, no. 1 (2021): 1-11.

[66] Mottaghi, Vahid, Mahdi Esmaeili, Ghasem Ali Bazaee, and Mohammadali Afshar Kazemi. "A decision-making system for detecting fake persian news by improving deep learning algorithms{case study of Covid-19 news." Journal of applied research on industrial engineering 8, no. Special Issue (2021): 1-17.

[67] Sadr, Mohammad Mohsen, Afshin Mousavi Chelak, Soraya Ziaei, and Jafar Tanha. "A predictive model based on machine learning methods to recognize fake persian news on twitter." International Journal of Nonlinear Analysis and Applications 11 (2020): 119-128.

[68] Sadr, Mohammad Mohsen. "The Use of LSTM Neural Network to Detect Fake News on Persian Twitter." Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12, no. 11 (2021): 6658-6668.

[69] Jahanbakhsh-Nagadeh, Zoleikha, Mohammad-Reza Feizi-Derakhshi, and Arash Sharifi. "A semi-supervised model for Persian rumor verification based on content information." Multimedia Tools and Applications 80, no. 28 (2021): 35267-35295.

[70] Jahanbakhsh-Nagadeh, Zoleikha, Mohammad-Reza Feizi-Derakhshi, and Arash Sharifi. "A Deep Content-Based Model for Persian Rumor Verification." Transactions on Asian and Low-Resource Language Information Processing 21, no. 1 (2021): 1-29.19

[71] Thoyyibah, Luthfiyatun. "Presupposition Triggers-a Comparative Analysis Between Oral News and Written Online News Discourse." JALL (Journal of Applied Linguistics and Literacy) 1, no. 2 (2017): 10-23.

[72] Cohen, Jacob. "A coefficient of agreement for nominal scales." Educational and psychological measurement 20, no. 1 (1960): 37-46.

[73] Krippendorff , Klaus. "Validity in content analysis." (1980): 69.

[74] 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).

[75] Farahani, Mehrdad, Mohammad Gharachorloo, Marzieh Farahani, and Mohammad Manthouri. "Parsbert: Transformer-based model for persian language understanding." Neural Processing Letters 53, no. 6 (2021): 3831-3847.

[76] 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).

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

[78] 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).

[79] Tenney, Ian, Patrick Xia, Berlin Chen, Alex Wang, Adam Poliak, R. Thomas McCoy, Najoung Kim et al. "What do you learn from context? probing for sentence structure in contextualized word representations." arXiv preprint arXiv:1905.06316 (2019).

[80] Hewitt, John, and Christopher D. Manning. "A structural probe for finding syntax in word representations". In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pp. 4129-4138. (2019).

[81] Ehsan Doostmohammadi, Minoo Nassajian, and Adel Rahimi. "Persian Ezafe Recognition Using Transformers and Its Role in Part-Of-Speech Tagging". In Findings of the Association for Computational Linguistics: EMNLP 2020, pages 961{971, Online. Association for Computational Linguistics. (2020).