NSURL-2021 Task 1: Semantic Relation Extraction in Persian

Nasrin Taghizadeh
University of Tehran
Tehran, Iran
nsr.taghizadeh@ut.ac.ir

Ali Ebrahimi
University of Tehran
Tehran, Iran
ali96ebrahimi@ut.ac.ir

Heshaam Faili
University of Tehran
Tehran, Iran
hfaili@ut.ac.ir

Abstract

Semantic Relation Extraction aims to identify whether a semantic relation of pre-defined types is held between two entities in a text. Relation extraction is a preliminary task in many applications such as knowledge base construction and information retrieval. To investigate the challenges and opportunities of relation extraction in Persian, we run a shared task as part of the second workshop on NLP Solutions for Under-Resourced Languages (NSURL 2021). This paper presents the approaches of the participating teams, their results, and the finding of the shared task. The data set prepared for this task is made publicly available[1] to support further researches on Persian relation extraction.

1 Introduction

The process of extracting structured information from unstructured text, known as information extraction, mostly consists of finding named entities (Taghizadeh et al., 2019), linking entities together, and extracting relations between them. Relation Extraction (RE) is a key component for building knowledge graphs, and it is of crucial significance to NLP applications such as structured search, question answering, and summarization.

RE is a well-studied task in English (Geng et al., 2020), Arabic (Taghizadeh et al., 2018) and Chinese (Li et al., 2019), regarding data sets of ACE, SemEval, TACRED, etc. However, due to the lack of public annotated corpora, the task is not highly examined in low-resource languages. Therefore, NSURL-2021 shared task 1 focuses on the relation extraction in Persian. The goal of the task is to specify whether a relationship exists between two entities in a Persian sentence, given a pre-defined set of semantic relations.

[1] https://github.com/nasrin-taghizadeh/NSURL-Persian-RelationExtraction

SemEval-2010 task 8 data set (Hendrickx et al., 2010) is de facto standard for RE. There is a machine-translated version of this data set in Persian, that was post-edited by humans, called PERLEX (Asgari-Bidhendi et al., 2020). PERLEX was used for training RE systems in Persian by running some of the state-of-art methods. Although this data set facilitates studying the task of RE in Persian, there is still a high need for an annotated data set developed from scratch, derived from Persian corpus, and reflects the common entities and new named entities appearing in Persian articles, news, social media, etc. Therefore, we prepared a data set of 1500 instances annotated with the semantic relations to be used as the test data of the shared task.

This paper presents a brief description of the participating teams, their approaches, the results, and the finding of the shared task. All solutions are based on the pre-trained language models (Devlin et al., 2018; Farahani et al., 2020), which are fined-tuned for RE. Proposed approaches differ in pre-processing steps, using syntactic features, and the architecture of deep models. The best F_1 score was obtained by an adaptation of an existing method, RIFRE (Zhao et al., 2021) on the Persian data set. Although, RIFRE obtained 91.3% of F_1 on SemEval 2010-task 8 data set, its score on the test set of PERLEX and test set of the shared task is 83.82% and 67.67%, respectively. Analysis of the results shows that new entities, misleading keywords, and complex grammatical structures are some reasons for the drop of the performance.

The rest of this paper is organized as follows: In Section 2, the definition of the shared task is presented. Section 3 contains an overview of the related works. Next, Section 4 describes the data set of the shared task. Section 5 includes the proposed solutions, their scores, and analytical results. Finally, Section 6 presents the conclusion remarks.

Table 1: Relation types of SemEval 2010- task 8 dataset (Hendrickx et al., 2010).

Relation Type	Definition
Cause-Effect(X, Y)	X is the cause of Y, or that X causes/makes/produces/emits/... Y.
Instrument-Agency(X, Y)	X is the instrument (tool) of Y or, equivalently, that Y uses X.
Product-Producer(X, Y)	X is a product of Y, or Y produces X.
Content-Container(X, Y)	X is or was (usually temporarily) stored or carried inside Y.
Entity-Origin(X, Y)	Y is the origin of an entity X (rather than its location), and X is coming or derived from Y.
Entity-Destination(X, Y)	Y is the destination of X in the sense of X moving (in a physical or abstract sense) toward Y.
Component-Whole(X,Y)	X has a functional relation with Y and X has an operating or usable purpose within Y.
Member-Collection(X, Y)	X is a member of Y.
Message-Topic(X, Y)	X is a communicative message containing information about Y.

2 Background

Persian is among the low-resource languages which suffer from lack of annotated data and preprocessing tools. However, language-specific features of Persian motivates researchers to develop customized machine learning methods. Therefore, it is crucial to create annotated data sets for different NLP tasks in Persian.

Given two entities in a text, the task is to predict the type of semantic relation between them, given a pre-defined set of relation types. Two entity mentions are tagged with e_1 and e_2 in the sentence. Each entity is a span over the sentence. Entities don't have a specific type and the numbering simply reflects the order of mentions in the sentence. The relation types of the shared task include 9 bi-directional relations defined in SemEval 2010-task 8, which are presented in Table 1. We defined two sub-tasks:

- Sub-Task A: Mono-Lingual Relation Extraction: In this subtask, the training data is in Persian. The aim is to use this data set for training.
- Sub-Task B: Bi-Lingual English-Persian Relation Extraction: In this subtask, the training data is a parallel English-Persian data set. The aim is to employ the bi-lingual data to train the model.

The prominent approach for both sub-tasks is to formulate them as a classification problem, however, the learning methods such as distant supervision, and bootstrapping are also applicable.

3 Related Works

Relation extraction has been extensively studied and a broad range of semantic relations has been examined by different researchers. ACE released a series of data sets in which the relations within the family, organization, society, etc. are mostly considered (Walker et al., 2005). SNPPhenA (Bokharaeian et al., 2017) considered the biological entities and relationships.

Since the importance of the RE, several shared tasks were held in different languages. Recently, SemEval-2020 Task 6 (DeftEval) (Spala et al., 2020) considered the problem of definition extraction, in which three subtasks are defined, one of them is to extract relation between terms and definitions. SemEval-2018 task 7 (Gábor et al., 2018) focused on relation extraction and classification in scientific paper abstracts, to extract specialized knowledge from domain corpora. In contrast, SemEval-2018 task 10 (Krebs et al., 2018) examined the task of identifying semantic difference which is a ternary relation between two concepts (e.g. apple, banana) and a discriminative attribute (e.g. red) that characterizes the first concept but not the other. WNUT-2020 Task 1 considered extracting entities and relations from wet-lab protocols. Wet-lab protocols consist of the guidelines from different lab procedures which involve chemicals, drugs, or other materials in liquid solutions or volatile phases (Tabassum et al., 2020).

There are a huge amount of researches on relation extraction. Recent methods are mainly based on the pre-trained language models such as BERT (Devlin et al., 2018), which are used to make a representation of samples with the same relation to be close to the representation of the corresponding relation in an embedding space. Cohen et al. (2020) proposed to utilize span-predictions models as used in question-answering models, by creating some questions based on sentences, then trying to find relations based on answers to these questions.

Graph neural networks have been employed to update sentence representation by message passing in the network to find a suitable relation for entities (Zhao et al., 2021, 2019). Peters et al. (2019) used a knowledge graph to enhance the representations of the words.

Many researchers showed that the syntactic features of the sentence are highly informative for the task of RE. Veyseh et al. (2020) utilized Ordered-Neuron Long-Short Term Memory Networks (ON-LSTM) to infer the model-based importance scores for RE for every word in the sentences that are then regulated to be consistent with the syntax-based scores to enable syntactic information injection. Tao et al. (2019) combined syntactic indicator and sequential context for relation prediction.

Since the lack of labeled data in many languages, multi-lingual and cross-lingual methods were proposed to benefit from the labeled data of high-resource languages in low-source languages. In this regard, Generative Adversarial Network (GAN) is used to transfer feature representations from one language with rich annotated data to another language with few annotated data (Zou et al., 2018). Taghizadeh et al. (2022) presented two deep CNN networks to employ syntactic features of the shortest dependency path between entities based on the Universal Dependencies.

4 Annotated Corpus

In this section, the data sets used for the development and evaluation of Persian RE systems are described.

4.1 Training and Development Data

The data set that used in the development phase is PERLEX, which is the translation of the SemEval-2010 task 8 data set. This data set has been already split into train and test with 8000 and 2717 samples, respectively. The test part can be used as the development set, or both parts can be combined and then divided randomly into the training and development sets.

4.2 Test Data

We have developed a data set of 1500 sentences annotated with two entities and the relationship held between them. Regarding language models such as BERT (Devlin et al., 2018), which improves the task of natural language understanding, some limitations of the old data sets like SemEval-2010 task 8 can be released in new data sets. Specifically, in the SemEval data set, entities are base Noun Phrases (NP) whose head is a common noun. We take into account 1) complex NPs (those NP with attached prepositional phrases), 2) nouns within verbal phrases, and 3) named entities in few instances, in addition to the base NPs. Moreover, in some instances, two entities are not in one sentence rather in two consecutive sentences. This data set also contains informal sentences. Table 3 shows some examples. Similar to the SemEval data set, we do not annotate examples whose interpretation relies on the discourse knowledge, and sentences with negation (e.g. no, not) whose scope contains the relation.

Table 2: Distribution of the task evaluation set in different semantic classes.

Class	(e1, e2)	(e2, e1)	Total
Cause–Effect	107	46	153
Component–Whole	86	45	131
Content–Container	62	51	113
Entity–Destination	137	20	157
Entity–Origin	108	30	138
Instrument–Agency	48	69	117
Member–Collection	92	48	140
Message–Topic	98	48	146
Product-Producer	80	90	170
Other	235		235
Total			1500

In the process of making the test set of the shared task, first, we collected a corpus of 50K sentences from the Virgool website. Virgool is a social network for sharing Persian articles [2]. This corpus was pre-processed, tokenized, and annotated by Part Of Speech (POS) tags. All nouns were considered as potential entities whose borders were revised later by human annotators. Next, we trained a state-of-the-art method using the PERLEX data set, to automatically annotate the relation held between every pair of entities in the sentences. At the next step, two human annotators corrected the automatic labels based on the annotation guideline of SemEval 2010- task 8. Since the semantic relations are language-independent, the English guideline is also useful for annotating Persian text. Finally, after several revisions of annotations, 1500 samples were selected. Table 2 shows the distribution of this data in different classes.

The annotators faced some challenges during the annotation of semantic relations. One chal-

[2] https://virgool.io/

Table 3: Examples of entities in test set of the shared task.

Entity	**English Equivalent**	**Persian Example**
complex NP	Even ⟨those⟩$_{e1}$ whose job is not subject to Corona's restrictions suffer from the economic impact of this ⟨epidemic⟩$_{e2}$.	حتی ⟨کسانی⟩$_{e1}$ که شغلشان مشمول محدودیتهای کرونایی نمی‌شود، از تاثیر اقتصادی این ⟨بیماری⟩$_{e2}$ همه‌گیر رنج می‌برند.
noun in VP	Sometimes ⟨exam pressure⟩$_{e1}$ can make you ⟨scared⟩$_{e2}$.	گاهی اوقات، ⟨فشار کنکور⟩$_{e1}$ می‌تواند شما را دچار ⟨وحشت⟩$_{e2}$ کند.
Named Entities	⟨Nazanin⟩$_{e1}$ is the only daughter in the ⟨family⟩$_{e2}$.	⟨نازنین⟩$_{e1}$ تنها دختر ⟨خانواده⟩$_{e2}$ است.
entities in two sentences	The height of this ⟨waterfall⟩$_{e1}$ is about 7 meters and it falls down from a ⟨rock wall⟩$_{e2}$.	ارتفاع این ⟨آبشار⟩$_{e1}$ هفت متر است و از یک ⟨دیواره صخره‌ای⟩$_{e2}$ به پایین می‌ریزد.
informal words	I can say that the first week of taking the ⟨medication⟩$_{e1}$ I was just ⟨asleep⟩$_{e2}$.	به جرات می‌تونم بگم که هفته اول مصرف ⟨داروها⟩$_{e1}$ فقط ⟨خواب⟩$_{e2}$ بودم.

lenge relates to the confusion of classes. For example, the relationship between entities in the following sentence may be confused among `Component-Whole`, `Content-Container`, and `Entity-Origin`:

⟨پرتقال⟩$_{e1}$ و گوجه فرنگی از منابع ⟨ویتامین سی⟩$_{e2}$ هستند.

⟨Orange⟩$_{e1}$ and tomato are the sources of ⟨vitamin C⟩$_{e2}$.

Considering the guideline of the shared task, `Component-Whole` shows the functional relationship between two entities, while `Content-Container` means that one entity is stored or carried inside another one. Therefore, `Entity-Origin` is the true label, which means that one entity is coming or derived from another one.

5 Experiments

In this section, we describe the participating teams, and then their results on the test data of the shared task. Finally, the analytical findings of the shared task are presented.

5.1 Participating Teams

The shared task was managed using the CodaLab competition platform[3] for result submission. A total of 4 systems has been submitted for sub-task A and no system for sub-task B. In the following, we describe the methodologies used by them.

HooshYar This team presented two methods for Persian RE. In both methods, they utilized the pre-trained language model of ParsBERT (Farahani et al., 2020) and fine-tuned it on the task of RE.

- In the first method, U-BERT, they attended to the class distribution of data and tried to improve the accuracy of the model using over-sampling of the instances of smaller classes. In addition, based on the fact that `Other` class contains many samples with diverse relations beyond the nine desired classes, they employed the Pairwise ranking loss function.
- In the second method, T-BERT, they focused on the syntactic features of the sentence. Many researchers used the shortest dependency path between two entities in the dependency tree of the sentence to recognize the relation held between them. Therefore, syntactic features inspire the use of a new embedding layer at the input of the BERT network. In this step, the vector for each word is reinforced with POS Tag and dependency tree tag. They used available tools in the Persian language to extract POS and dependency tree tags of the sentences. In the last layer of their network, they used the vector of average entity words in addition to the CLS token for classification.

SBU-NLP This team performed some pre-processing steps on PERLEX. Since it is a semi-automatic translated data set, they removed those samples with more than one entity marker (<e1> and </e1>), or unclear translation. Moreover, they used data augmentation techniques and back-translation methods to increase training data size. They inspired the R-BERT model (Wu and He, 2019) and examined several changes in the architectures of this network to improve model accuracy including 1) averaging both of the three final segments in the R-BERT rather than a concatenation of them, 2) concatenation of all of the tokens in the entities rather than average them, 3) using the last (first) token instead of average all of the to-

[3]https://competitions.codalab.org/competitions/31979

kens in the entities, and 4) using the Multilingual BERT (mBERT) (Devlin et al., 2018) and ParsBERT (Farahani et al., 2020) to reaching the best decision.

Customizing the available methods One of the participating teams adapted the method proposed by We and He (2019), called R-BERT. They used ParsBERT (Farahani et al., 2020), a pre-trained language model for Persian, and set the parameters of the model to the best-fit values on the PERLEX data set. Therefore, we refer to this method as R-BERT+`ParsBERT`.

5.2 Results

Table 4 shows a summary of results for the participating teams. We reported the F_1 score for every relation in addition to the macro-average F_1 considering the direction of the relations. The first part of Table 4 contains the evaluation results on the official test set of the shared task, where all data of PERLEX (10,717 samples) can be used for training the systems. The second part of Table 4 presents the F_1 scores of the same methods when trained with the training part of PERLEX (8000 samples) and evaluated by the test part of PERLEX (2717 samples).

For better comparison, we also reported the result of the state-of-the-art method of Zhao et al. (2021), named RIFRE. They used graph neural networks and modeled relations and words as nodes on the graph and fuse the two types of semantic nodes by the message passing mechanism iteratively to obtain nodes representation that is more suitable for the RE task. We used ParsBERT as the encoder layer of the network and fine-tuned it on PERLEX. This method obtained the top rank on the English data set of SemEval 2010-task 8.

As Table 4 shows, the F_1 scores on shared task data are much lower than PERLEX test data for all methods. Among five methods, the state-of-the-art methods of RIFRE+`ParsBERT` obtained the highest F_1 scores on both test data of the shared task, 67.67% F_1, and PERLEX, 83.82% F_1; while this method obtained 91.3% score of F_1 on English equivalent data set (SemEval 2010-task 8).

Due to the several improvements over R-BERT+`ParsBERT` made by the method proposed by Moein Salimi (Salimi Sartakhti et al., 2021), this method outperformed R-BERT+`ParsBERT` on PERLEX test data, however, it obtained a lower F_1 score on the test set of the shared task.

5.3 Analysis

Although the state-of-the-art RE methods obtained more than 90% of F_1 score on SemEval 2010-task 8 data set (Cohen et al., 2020; Zhao et al., 2021), their performances drop in Persian. We investigate the impact of new entities, misleading keywords, and complex grammatical structures.

New Entities Comparing the F_1 scores which are obtained on the test data of PERLEX with those reported on the test data of the shared task in Table 4 reveals that there is a drop in results. One reason is that the shared task test data contains the new entities that do not appear in PERLEX. Statistics show about 70% of entities are new. Moreover, the shared task test data contains some samples that flout the guidelines of SemEval 2010-task 8 regarding the locality of entities, nominal expression, etc., as depicted in Table 3.

Misleading Keywords Have a deeper look at the performance of the models, several keywords specify each class. For example, `Cause-Effect` is usually specified by words such as "cause/ caused by/ result/ generate/ triggered/ due/ effect" (Taghizadeh and Faili, 2021). There are similar keywords in Persian such as "تاثیر، باعث، سبب، موجب". However, some sentences have these keywords but lack the corresponding relation:

⟨سالمندان⟩$_{e1}$ به دلیل تاثیر این دارو در ⟨خونریزی⟩$_{e2}$ و عدم هماهنگی ، باید در مصرف آن احتیاط کنند.

⟨The elderly⟩$_{e1}$ should avoid taking this drug due to its effect on ⟨bleeding⟩$_{e2}$ and lack of coordination.

The relation of this example is `Other`, not `Cause-Effect`. We intentionally gathered such examples in the test data of the shared task. Most models fail to recognize the true relation of these samples. Therefore these models mainly memorize the keywords surrounding the entities rather than understanding the semantic relations between them.

On the other hand, some relation instances lack any keywords, such as the following example, where a `Cause-Effect` relation is held between entities:

تنها چیزی که می تواند وضعیت جاری را دچار تحول کند و ⟨نیروی محرکه⟩$_{e1}$ است ⟨تجارت⟩$_{e2}$ است.

The only thing that can change the current situation and act as ⟨propulsion⟩$_{e1}$, is ⟨trading⟩$_{e2}$.

Table 4: Results of the participating teams against the state-of-the-are approaches for mono-lingual RE (Sub-Task A).

	Cause–Effect	Component–Whole	Content–Container	Entity–Destination	Entity–Origin	Instrument–Agency	Member–Collection	Message–Topic	Product–Producer	F1
Official test set of the shared task										
T-Bert (Jafari et al., 2021)	56.74	56.05	49.14	71.43	56.93	59.93	43.87	60.95	63.32	57.60
U-BERT (Jafari et al., 2021)	58.33	55.75	50.91	69.48	59.06	66.92	47.23	65.93	61.35	59.44
SBU-NLP (Salimi Sartakhti et al., 2021)	61.70	66.44	59.26	76.01	58.04	75.54	32.85	76.06	76.13	64.67
R-BERT (Wu and He, 2019) + ParsBERT	62.76	62.14	55.37	75.17	66.19	74.72	50.66	73.00	79.13	66.57
RIFRE (Zhao et al., 2021) + ParsBERT	72.11	59.93	51.25	76.77	71.79	74.36	53.95	70.73	78.15	67.67
Test set of PERLEX										
T-Bert (Jafari et al., 2021)	88.11	74.14	80.00	84.81	75.39	61.05	72.53	81.80	74.90	76.97
U-BERT (Jafari et al., 2021)	88.72	74.41	82.38	85.01	76.98	72.85	73.57	78.57	77.02	78.83
R-BERT (Wu and He, 2019) + ParsBERT	87.91	73.29	79.81	85.97	76.60	74.07	73.89	83.11	77.35	79.11
SBU-NLP (Salimi Sartakhti et al., 2021)	89.37	77.45	82.13	88.58	79.84	76.07	76.60	85.92	79.91	81.76
RIFRE (Zhao et al., 2021) + ParsBERT	93.07	80.54	80.11	85.76	81.92	80.39	85.40	90.41	76.79	83.82

Complex Syntactic Structures Many researchers used the shortest dependency path between entities to detect their relation type. However, when two entities are in separate sentences or complex structures, syntax-based methods usually fail to predict the correct relation, mainly due to the low accuracy of the dependency parser.

6 Conclusion

In this paper, we described the Persian relation extraction shared task that was organized in NSURL-2021. We developed test data that is publicly available. This Persian corpus was developed from scratch, against PERLEX data set that is a semi-automatic translated data. This corpus facilitates further researches on Persian RE.

References

Majid Asgari-Bidhendi, Mehrdad Nasser, Behrooz Janfada, and Behrouz Minaei-Bidgoli. 2020. Perlex: A bilingual persian-english gold dataset for relation extraction. *arXiv preprint arXiv:2005.06588*.

Behrouz Bokharaeian, Alberto Diaz, Nasrin Taghizadeh, Hamidreza Chitsaz, and Ramyar Chavoshinejad. 2017. SNPPhenA: a corpus for extracting ranked associations of single-nucleotide polymorphisms and phenotypes from literature. *Journal of biomedical semantics*, 8(1):14.

Amir DN Cohen, Shachar Rosenman, and Yoav Goldberg. 2020. Relation classification as two-way span-prediction. *arXiv preprint arXiv:2010.04829*.

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

Mehrdad Farahani, Mohammad Gharachorloo, Marzieh Farahani, and Mohammad Manthouri. 2020. Parsbert: Transformer-based model for persian language understanding. *arXiv preprint arXiv:2005.12515*.

Kata Gábor, Davide Buscaldi, Anne-Kathrin Schumann, Behrang QasemiZadeh, Haïfa Zargayouna, and Thierry Charnois. 2018. SemEval-2018 task 7: Semantic relation extraction and classification in scientific papers. In *Proceedings of The 12th International Workshop on Semantic Evaluation*, pages 679–688, New Orleans, Louisiana. Association for Computational Linguistics.

ZhiQiang Geng, GuoFei Chen, YongMing Han, Gang Lu, and Fang Li. 2020. Semantic relation extraction using sequential and tree-structured lstm with attention. *Information Sciences*, 509:183–192.

Iris Hendrickx, Su Nam Kim, Zornitsa Kozareva, Preslav Nakov, Diarmuid Ó Séaghdha, Sebastian Padó, Marco Pennacchiotti, Lorenza Romano, and Stan Szpakowicz. 2010. Semeval-2010 task 8: Multi-way classification of semantic relations between pairs of nominals. In *Proceedings of the Workshop on Semantic Evaluations: Recent Achievements and Future Directions*, pages 94–99s.

Mohammad Mahdi Jafari, Somayyeh Behmanesh, Alireza Talebpour, and Ali Nadian Ghomsheh. 2021. Improving pre-trained language model for relation extraction using syntactic information in persian. In *Proceedings of The Second International Workshop on NLP Solutions for Under Resourced Languages (NSURL 2021) co-located with ICNLSP 2021 - Short Papers*, Trento, Italy.

Alicia Krebs, Alessandro Lenci, and Denis Paperno. 2018. SemEval-2018 task 10: Capturing discriminative attributes. In *Proceedings of The 12th International Workshop on Semantic Evaluation*, pages 732–740, New Orleans, Louisiana. Association for Computational Linguistics.

Ziran Li, Ning Ding, Zhiyuan Liu, Haitao Zheng, and Ying Shen. 2019. Chinese relation extraction with multi-grained information and external linguistic knowledge. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4377–4386.

Matthew E Peters, Mark Neumann, Robert L Logan IV, Roy Schwartz, Vidur Joshi, Sameer Singh, and Noah A Smith. 2019. Knowledge enhanced contextual word representations. *arXiv preprint arXiv:1909.04164*.

Moein Salimi Sartakhti, Romina Etezadi, and Mehrnoosh Shamsfard. 2021. Persian relation extraction using ParsBERT on the PERLEX dataset. In *Proceedings of The Second International Workshop on NLP Solutions for Under Resourced Languages (NSURL 2021) co-located with ICNLSP 2021 - Short Papers*, Trento, Italy.

Sasha Spala, Nicholas Miller, Franck Dernoncourt, and Carl Dockhorn. 2020. SemEval-2020 task 6: Definition extraction from free text with the DEFT corpus. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*, pages 336–345, Barcelona (online). International Committee for Computational Linguistics.

Jeniya Tabassum, Wei Xu, and Alan Ritter. 2020. WNUT-2020 task 1 overview: Extracting entities and relations from wet lab protocols. In *Proceedings of the Sixth Workshop on Noisy User-generated Text (W-NUT 2020)*, pages 260–267, Online. Association for Computational Linguistics.

Nasrin Taghizadeh, Zeinab Borhanifard, Melika Golestani Pour, Mojgan Farhoodi, Maryam Mahmoudi, Masoumeh Azimzadeh, and Hesham Faili. 2019. NSURL-2019 task 7: Named entity recognition for Farsi. In *Proceedings of The First International Workshop on NLP Solutions for Under Resourced Languages (NSURL 2019) co-located with ICNLSP 2019 - Short Papers*, pages 9–15, Trento, Italy. Association for Computational Linguistics.

Nasrin Taghizadeh and Heshaam Faili. 2021. Cross-lingual adaptation using universal dependencies. *Transactions on Asian and Low-Resource Language Information Processing*, 20(4):1–23.

Nasrin Taghizadeh and Heshaam Faili. 2022. Cross-lingual transfer learning for relation extraction using universal dependencies. *Computer Speech & Language*, 71:101265.

Nasrin Taghizadeh, Heshaam Faili, and Jalal Maleki. 2018. Cross-language learning for arabic relation extraction. *Procedia computer science*, 142:190–197.

Qiongxing Tao, Xiangfeng Luo, Hao Wang, and Richard Xu. 2019. Enhancing relation extraction using syntactic indicators and sentential contexts. In *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 1574–1580. IEEE.

Amir Pouran Ben Veyseh, Franck Dernoncourt, Dejing Dou, and Thien Huu Nguyen. 2020. Exploiting the syntax-model consistency for neural relation extraction. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 8021–8032.

Christopher Walker, Stephanie Strassel, Julie Medero, and Kazuaki Maeda. 2005. Ace 2005 multilingual training corpus-linguistic data consortium. *URL: https://catalog. ldc. upenn. edu/LDC2006T06*.

Shanchan Wu and Yifan He. 2019. Enriching pre-trained language model with entity information for relation classification. In *Proceedings of the 28th ACM international conference on information and knowledge management*, pages 2361–2364.

Kang Zhao, Hua Xu, Yue Cheng, Xiaoteng Li, and Kai Gao. 2021. Representation iterative fusion based on heterogeneous graph neural network for joint entity and relation extraction. *Knowledge-Based Systems*, page 106888.

Yi Zhao, Huaiyu Wan, Jianwei Gao, and Youfang Lin. 2019. Improving relation classification by entity pair graph. In *Asian Conference on Machine Learning*, pages 1156–1171. PMLR.

Bowei Zou, Zengzhuang Xu, Yu Hong, and Guodong Zhou. 2018. Adversarial feature adaptation for cross-lingual relation classification. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 437–448.