

Incorporating Type Information into Zero-Shot Relation Extraction

Cedric Möller^{1,*}, Ricardo Usbeck²

¹Universität Hamburg, Department of Informatics, Semantic Systems, Germany

²Leuphana Universität Lüneburg, Institute for Information Systems, Artificial Intelligence and Explainability, Germany

Abstract

The task of zero-shot relation extraction focuses on the extraction of relations not seen during training time. Commonly, additional information about the relation such as the relation name or a description of the relation is utilised. In this work, we analyze whether a relation extractor can benefit from the inclusion of fine-grained type information about the involved entities. This is based on the intuition that relation descriptions might contain ontological information on the domain and range of the entity types that are usually put into relation. For that, we follow a cross-encoding setup where we encode both, the entity information and relation information, as one sequence and learn to score the representation. We examine this method on several datasets and show that the inclusion of the fine-grained type information leads to an improvement in performance.

Keywords

Relation Extraction, Zero-shot, Entity types

1. Introduction

Identifying the relation that is expressed between entities is a very important subproblem of various downstream tasks. For instance, it is critical to handle semantic-web-related tasks such as knowledge graph question answering or knowledge graph population. Usually, it is assumed that the encountered relations are known before. Zero-shot relation extraction breaks with this assumption. During inference time, the goal is to extract entirely new relations not seen before during training time.

With the establishment of pre-trained models, this goal becomes achievable. Those models are trained on large corpora of textual data in an unsupervised way. In zero-shot relation extraction, one assumes that some information on the new relations is available. The simplest kind of information is a label describing the relation. But this only works if the relation label co-occurs with a similar context as encountered during the training of the pre-trained models. If this is not the case, using additional information such as a description of the relation is necessary.

In this work, we analyse the impact of combining fine-grained type information and the relation description on the relation extraction performance. This is based on the assumption that the descriptions contain valuable information on the types of the involved entities. For example, the description of the relation `director` states `director(s) of film, TV-series, stageplay, video game or similar`. Therefore it is clear, that the relation should not be used when talking about board members of a company, also sometimes referred to as directors. We incorporate fine-grained type information extracted from **Wikidata** together with the relation descriptions in the relation extraction process.¹

The contributions are:

- Zero-shot relation extraction model using fine-grained type information and relation descriptions
- Ablation study on the impact of fine-grained type information and relation descriptions on the performance

TEXT2KG 2024: Third International Workshop on Knowledge Graph Generation from Text, May 26-30, 2024, co-located with Extended Semantic Web Conference (ESWC), Hersonissos, Greece

*Corresponding author.

✉ cedric.moeller@uni-hamburg.de (C. Möller); ricardo.usbeck@leuphana.de (R. Usbeck)

🆔 0000-0001-6700-3482 (C. Möller); 0000-0002-0191-7211 (R. Usbeck)

© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

¹Code/Data available at: <https://github.com/semantic-systems/zero-shot-re>

2. Methodology

2.1. Problem Definition

The problem of relation extraction can be defined as follows: Given an input text c , an annotated head h and tail entity e , identify the correct relation r as expressed in the text. Zero-shot relation extraction separates the set of relations encountered during training from the ones encountered during inference. Hence, during training time, the set of available relations is R_{train} , while during test time, the set is R_{test} . It holds that $R_{\text{train}} \cap R_{\text{test}} = \emptyset$. Also, no annotated examples containing any relations in set R_{test} are available during inference time. Additional information defining the relation is available. We assume labels, descriptions and type information on entities to be available.

2.2. Method

To study the impact of fine-grained type information, we opt to extend a simple but powerful model introduced by Lan et al. [1]. Hence, we cross-encode the information of the text and the relation information in a single input. Different from their work, we do not solely rely on the relation label but also include the relation description. Additionally, we assume the existence of fine-grained types for both, the head and the tail relation, extracted using the P31 relation in Wikidata. We include the relation description under the assumption that it contains valuable ontological information referring to the fine-grained types of the considered entities. For example, for the relation `shipping port`, the description is

```
shipping port of the vessel (if different from
"ship registry"): For civilian ships, the
primary port from which the ship operates ...
```

We denote the types of the head entity by \mathcal{T}_h and the types of the tail entity by \mathcal{T}_t . Additionally, for each type of an entity, we extract the label describing the type (e.g., human for Q5). The input x to the model then consists of four different segments. The **first segment** describes the head entity:

```
Head Entity :  $\{l_h\}$  with Types :  $\{T_h\}$ 
```

and the **second segment** describes the tail entity:

```
Tail Entity :  $\{l_t\}$  with Types :  $\{T_t\}$ 
```

where l_{\square} denotes the label of the head entity h or tail entity t . T_{\square} is the concatenation of the labels of the types of the head or tail entity $T_{\square} = \bigoplus_{u \in \mathcal{T}_{\square}} l_u$.

The **third segment** gives information on the input text:

```
Context :  $\{c\}$ 
```

The **final segment** gives information on the relation:

```
 $\{l_r\}$  defined as  $\{d_r\}$ 
```

where l_r denotes the label of the relation r and d_r is the description of the relation r .

All segments are then combined into a single coherent text as follows:

```
[CLS] Given the Head Entity :  $\{l_h\}$  with
Types :  $\{T_h\}$ , Tail Entity :  $\{l_t\}$  with
Types :  $\{T_t\}$  and Context :  $\{c\}$ , the context
expresses the relation [SEP]  $\{l_r\}$  defined
as  $\{d_r\}$  [SEP]
```

The whole text is then fed into an encoder-only model $f(x)$ which returns a sequence of vectors $e_{[CLS]} \dots e_{[SEP]}$. The vector $e_{[CLS]}$ is then taken and fed a input to a linear layer which then returns a final score.

$$s_r = l(e_{[CLS]})$$

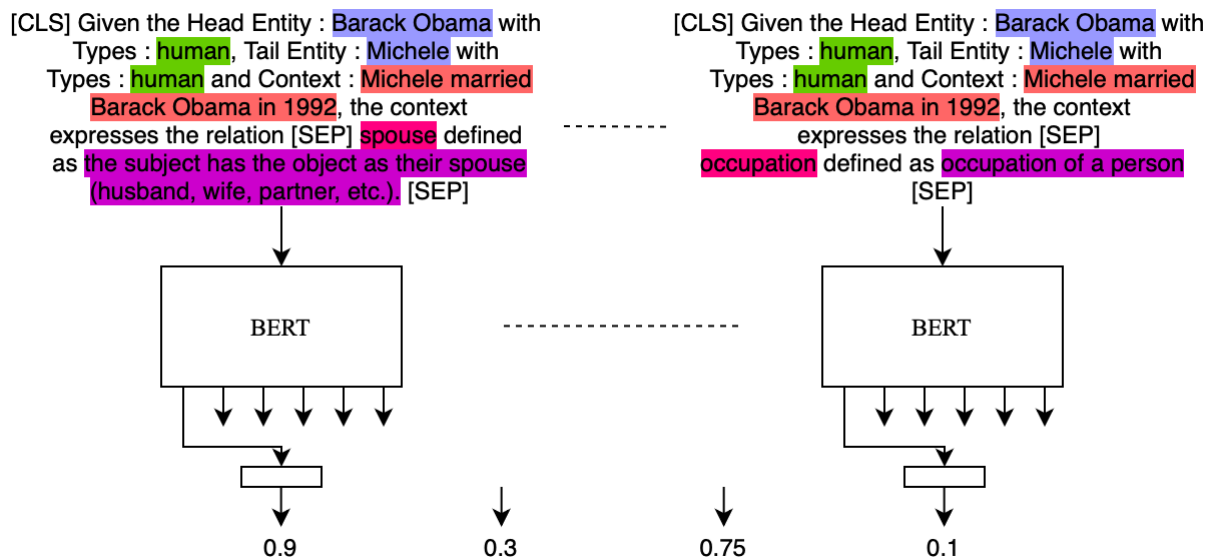


Figure 1: Model overview: Green specifies the types, blue the entities, orange the context, red the relation label and Purple the description of the relation.

This is done for each potential relation, which gives us $|R_{\text{test}}|$ scores. The highest score is taken as the predicted relation. All potential relations are known beforehand. During training, the model is optimized using cross-entropy loss. Each example contains a single positive relation. The model trains to differentiate it against other relations by comparing it against incorrect relations. For that, n other relations are sampled and used as negative examples.²

An overview of the model can be found in Figure 1.

3. Evaluation

We evaluate the model on two popular datasets, FewRel and Wiki-ZSL. Both datasets were annotated on Wikipedia article texts. FewRel is originally a few-shot relation extraction dataset annotated by Han et al. [2]. The dataset was modified for zero-shot purposes by Chia et al. [3]. They split the training, validation and test examples by their relations into disjoint sets. Wiki-ZSL is a zero-shot relation extraction dataset created by Chen et al. [4] based on the Wiki-KB [5]. As the entities and relations in both datasets are linked to Wikidata, we focus on it as the knowledge graph providing the fine-grained entity types.

In each dataset, the set of relations in the training and test dataset is disjoint and randomly assigned. Three different settings are examined per dataset. Each setting considers a different number of relations in the train, validation and test set. The number of relations in the validation/test set varies between $m = 5$, $m = 10$ and $m = 15$ relations. These relations are randomly picked and the remaining relations are assigned to the training set.

To handle the considerable noise induced by the random selection of the relations, the dataset for $m = 5$, $m = 10$ and $m = 15$ were randomly split into train, validation and test sets for five times. A method is evaluated on each split and the results are averaged.

As metrics, precision, recall and F1 are calculated. All metrics are computed in a macro setting which means that for each relation the precision, recall and F1 are calculated and then averaged over all relations.

We compare our method, called TMC-BERT, against several methods: CIM [6] solves the task as a textual entailment problem where the relation descriptions and the input sentence are given to a Natural Language Inference model to classify whether the input sentence entails the relation description. This

²In our experiments we set $n = 5$.

is done for all potential relations and the highest scoring is taken. ZS-BERT [7] encodes the input sentence as well as the relation descriptions into a dense vector space. An nearest neighbor search is conducted over all the encodings of the relation descriptions given the input sentence. The closest relation encoding is the final relation. Tran et al. (2022) [8] again encode the input sentence and relation descriptions into a dense vector space. They additionally employ a contrastive-learning inspired loss on the input sentence and relation encodings. The final scoring is achieved by concatenating the relation encoding and the sentence encoding and feeding it into a linear layer. RE-Matching [9] encodes the input sentence and relation descriptions as well but uses feature distillation to calculate a similarity score based on more fine-grained feature interactions. RelationPrompt [3] relies on a generative model to generate synthetic data as additional training samples. At the same time, the generative model is also used to generate a relation given the sentence and the two entities as input. We compare against the model with (RelationPrompt) and without (RelationPrompt NG) synthetic training data. MC-BERT [1] models the relation extraction similar to us as a multiple-choice problem where the input sentence and the relation label are rearranged together in a natural sentence, encoded and scored. DSP-ZRSC [10] solves the problem via Discriminative Soft Prompting where the input text, the entities and all relation labels are concatenated, fed into a prompt discriminative language model and each relation label is scored. Tran et al. (2023) [11] solve it as a representation learning problem and introduce a second loss term incorporating the degree of correlation between sentences and relations.

BERT-base-case was used as the model to stay comparable to MC-BERT. The model was fine-tuned on two NVIDIA A6000s with a batch size of 48 and a learning rate of $5e - 5$.

3.1. Results

As can be seen in Table 1, the incorporation of type-related information leads to a large increase in performance on several datasets in comparison to regular MC-BERT. On Wiki-ZSL, the performance increases vary between 6 and nearly 8 F1 points. The type-related information has a great impact on, both, recall and precision. On FewRel, the performance increases when considering 5 or 15 unseen relations. However, the performance increases are less pronounced. In comparison to the current SOTA method by Tran et al. [11], TMC-BERT considerably surpasses its performance when confronted with 15 unseen relations. This is the most complex setting as much fewer examples and relations are available during training while more potential relations are encountered during inference. Here, the additional type information helps a lot. Furthermore, the inclusion of fine-grained type information is orthogonal to the properties of the method by Tran et al. [11]. Their method could benefit from it as well.

3.2. Ablation study

To examine what changes lead to the large increase in performance, we conducted an ablation study on the incorporation of different kinds of information. Here, we differentiated between three cases:

1. TMC-BERT
2. TMC having the types of the subject and object entity removed (TMC-BERT w/o types)
3. TMC having the description of the relation removed (TMC-BERT w/o desc.)

As can be seen in Table 2, the addition of the relation description alone was the least beneficial type of information. Adding information on relation types leads to a larger improvement, probably as the pre-trained model already associates specific types with certain relation labels. Finally, the ablation study shows that the relation description and fine-grained entity type information complement each other, as using each separately does not lead to as large a decrease in performance as using them together.

3.3. Entity Linking impact

As it is not realistic that fine-grained type information is always available, we also evaluate the model when identifying entity types using an entity linker (EL). For that, we train the model with known

m	Model	Wiki-ZSL			FewRel		
		P	R	F1	P	R	F1
5	CIM	49.63	48.81	49.22	58.05	61.92	59.92
	ZS-BERT	71.54	72.39	71.96	76.96	78.86	77.90
	Tran et al. (2022)	87.48	77.50	82.19	87.11	86.29	86.69
	RelationPrompt NG	51.78	46.76	48.93	72.36	58.61	64.57
	RelationPrompt	70.66	83.75	76.63	90.15	88.50	89.30
	RE-Matching	78.19	78.41	78.30	92.82	92.34	92.58
	DSP-ZRSC	<u>94.1</u>	77.1	84.8	93.4	92.5	92.9
	Tran et al. (2023)	94.50	96.48	95.46	96.36	96.68	96.51
	MC-BERT	80.28	84.03	82.11	90.82	90.13	90.47
TMC-BERT	90.11	<u>87.89</u>	<u>88.92</u>	<u>93.94</u>	<u>93.30</u>	<u>93.62</u>	
10	CIM	46.54	47.90	45.57	47.39	49.11	48.23
	ZS-BERT	60.51	60.98	60.74	56.92	57.59	57.25
	Tran et al. (2022)	71.59	64.69	67.94	64.41	62.61	63.50
	RelationPrompt NG	54.87	36.52	43.80	66.47	48.28	55.61
	RelationPrompt	68.51	74.76	71.50	80.33	79.62	79.96
	RE-Matching	74.39	73.54	73.96	83.21	82.64	82.93
	DSP-ZRSC	80.0	74.0	76.9	80.7	88.0	84.2
	Tran et al. (2023)	85.43	88.14	86.74	81.13	82.24	81.68
	MC-BERT	72.81	73.96	73.38	86.57	<u>85.27</u>	85.92
TMC-BERT	<u>81.21</u>	<u>81.27</u>	<u>81.23</u>	<u>84.42</u>	84.99	<u>85.68</u>	
15	CIM	29.17	30.58	29.86	31.83	33.06	32.43
	ZS-BERT	34.12	34.38	34.25	35.54	38.19	36.82
	Tran et al. (2022)	38.37	36.05	37.17	43.96	39.11	41.36
	RelationPrompt NG	54.45	29.43	37.45	66.49	40.05	49.38
	RelationPrompt	63.69	<u>67.93</u>	65.74	74.33	72.51	73.40
	RE-Matching	67.31	67.33	67.32	73.80	73.52	73.66
	DSP-ZRSC	<u>77.5</u>	64.4	<u>70.4</u>	82.9	78.1	<u>80.4</u>
	Tran et al. (2023)	64.68	65.01	65.30	66.44	69.29	67.82
	MC-BERT	65.71	67.11	66.40	80.71	<u>79.84</u>	80.27
TMC-BERT	73.62	74.07	73.77	<u>82.11</u>	79.93	81.00	

Table 1
Results on FewRel and Wiki-ZSL

m	Model	Wiki-ZSL			FewRel		
		P	R	F1	P	R	F1
5	TMC-BERT w/o desc.	85.56	84.07	84.74	93.96	93.26	93.61
	TMC-BERT w/o types	85.00	84.41	84.68	93.33	92.50	92.91
	TMC-BERT	90.11	87.89	88.92	93.94	93.30	93.62
10	TMC-BERT w/o desc.	77.26	78.16	77.70	85.24	83.29	84.25
	TMC-BERT w/o types	74.89	76.05	75.46	85.16	83.36	84.24
	TMC-BERT	81.21	81.27	81.23	84.42	84.99	85.68
15	TMC-BERT w/o desc.	72.33	71.16	71.73	79.22	76.46	79.79
	TMC-BERT w/o types	68.53	69.81	69.16	79.22	78.19	78.69
	TMC-BERT	73.62	74.07	73.77	82.11	79.93	81.00

Table 2
Ablation study of on FewRel and Wiki-ZSL

entity types but evaluate with the entity types as retrieved by an entity linker. As an entity linker, we use ReFinED [12]. As can be seen in Table 3, the performance diminishes when using types identified through entity linking. On Wiki-ZSL, the performance is still surpassing the existing SOTA results

at all times. On FewRel, the performance is still greater when only confronted with five relations but decreases more when having to predict 10 or 15 relations. One reason might be that the entity linking performance is lower on FewRel than on Wiki-ZSL.

m	Model	Wiki-ZSL			FewRel		
		P	R	F1	P	R	F1
5	TMC-BERT	90.11	87.89	88.92	93.94	93.30	93.62
	TMC-BERT + EL	88.44	87.07	87.73	93.94	93.35	93.64
10	TMC-BERT	81.21	81.27	81.23	84.42	84.99	85.68
	TMC-BERT + EL	81.16	81.22	81.18	84.88	83.43	84.14
15	TMC-BERT	73.62	74.07	73.77	82.11	79.93	81.00
	TMC-BERT + EL	73.53	73.96	73.67	80.87	78.74	79.78

Table 3
Results on FewRel and Wiki-ZSL when using an entity linker

3.4. Case study

Table 4 illustrates two instances where the inclusion of type information or relation descriptions proved beneficial. In the first case, specifying that MMORPG belongs to the video game genre facilitated the correct classification of the **genre** relation. In the second example, highlighting that bass is a voice type aligned the type label precisely with the **voice type** relation label. Additionally, the relation description directly addressed the voice type of bass.

Method	TMC-BERT	MC-BERT
Sentence	Gravity Corporation is a South Korean video game corporation primarily known for the development of the MMORPG Ragnarok Online. {MMORPG: video game genre; Ragnarok Online: video game}	
Classified Relation	genre	manufacturer
Description of Relation	creative work’s genre or an artist’s field of work	main use of the subject (includes current and former usage)
Sentence with entity types	Putnam Griswold (1875-1914) was an American opera singer (bass), born in Minneapolis, Minnesota. {Putnam Griswold: human, bass: voice type}	
Classified Relation	voice type	use
Description of Relation	person’s voice type. expected values: soprano, mezzo-soprano, contralto, countertenor, tenor, baritone, bass (and derivatives)	main use of the subject (includes current and former usage)

Table 4
Comparison of the performance of TMC-BERT and MC-BERT on two different examples. Ground-truth relations are shown in bold. The interacting entities and their types are shown in dictionaries following the sentences.

4. Related Work

Commonly, relation extraction is tackled as classification problem. Usually, the input text is encoded and a classification head is attached. To encode text, CNNs [13], RNNs [14] or transformers [15] are usually employed. Recently, pre-trained models have been fine-tuned on the relation extraction task. Due to the fixed classification head, such trained models are not flexible enough to handle new relations [16]. Hence,

when targeting zero-shot relation extraction other methods are necessary. Representation-learning-based methods [11, 8, 4, 9, 11] try to embed the textual information and the relational information in the same vector space. For that, relational information such as labels or descriptions are usually used to get a representation of the relation. The goal is to learn representations such that the representation of the true relation resides close to a representation of the text in the vector space while the representation of false relations is far away. Recently, generative language models have been increasingly utilized for the task [17, 18, 3, 10]. Here, the model is prompted with the input text as well as information on the potential relations. The model is then fine-tuned to either generate the relation as expressed in the input text or a full triple consisting of the two entities and relation. For example, Chen et al. [18] model it as solving a Masked Language Modelling problem. Also, generative models were applied to generate synthetic training data for relation extraction [3].

Type information was considered in previous works focusing on relation extraction but these works either used very broad types or did not tackle zero-shot relation extraction [19, 20, 21].

Some methods model the problem as a textual entailment problem [22, 23, 24]. Here, the idea is that a model that is pre-trained on the textual entailment task is directly applied to the relation extraction task. The assumption is that the model can identify whether the textual information entails the relation description.

The method by Lan et al [1] models relation classification as a multiple-choice problem where the text is encoded with relation information and a score is calculated. This is done for all relations and the relation with the highest score is taken. We extend this approach.

5. Conclusion and Future Work

In this work, we examined the impact of fine-grained type information on the zero-shot relation extraction problem. Different from past methods, we employed fine-grained type information as additional information and showed that combining this information with the description of the relation leads to a synergistic effect, improving the performance overall. We believe that this is the case because the description information provides valuable ontological information on the domain and range of a relation. This domain and range are then compared against the fine-grained type information of the interacting entities. Furthermore, we validated whether the increase in performance did indeed spring from the combination of type and relation description information which is indeed the case. Finally, we studied the impact of using an entity linker to retrieve the entity types. While it leads to a decrease, the performance often still surpasses the current SOTA in the most complex setting considerably.

In future works, we want to tackle multiple problems. First, it is not certain that one has access to fine-grained type information during inference. Therefore, we want to examine, whether the performance of a trained entity typer is sufficient to produce similar results. Secondly, the current architecture follows a cross-encoding approach. While this is not a problem when one encounters only a few relations during inference, in real-world use cases this is not typically the case. There are hundreds of potential different relations that could be encountered during inference. Cross-encoding the text with each one leads to a substantial computational effort. We want to examine whether the relation candidate generation module might also benefit from fine-grained type information. Also, the training process currently only trains the model by randomly sampling other relations. Choosing the relationships in a smarter way might lead to additional improvement. Finally, the impact of fine-grained entity types from other knowledge graphs needs to be evaluated.

Acknowledgments

This project was supported by the House of Computing and Data Science (HCDS) of the Hamburg University within the Cross-Disciplinary Lab programme.

References

- [1] Y. Lan, D. Li, Y. Zhang, H. Zhao, G. Zhao, Modeling zero-shot relation classification as a multiple-choice problem, in: International Joint Conference on Neural Networks, IJCNN 2023, Gold Coast, Australia, June 18-23, 2023, IEEE, 2023, pp. 1–8. doi:10.1109/IJCNN54540.2023.10191459.
- [2] X. Han, H. Zhu, P. Yu, Z. Wang, Y. Yao, Z. Liu, M. Sun, Fewrel: A large-scale supervised few-shot relation classification dataset with state-of-the-art evaluation, in: E. Riloff, D. Chiang, J. Hockenmaier, J. Tsujii (Eds.), Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Brussels, Belgium, October 31 - November 4, 2018, Association for Computational Linguistics, 2018, pp. 4803–4809. doi:10.18653/V1/D18-1514.
- [3] Y. K. Chia, L. Bing, S. Poria, L. Si, Relationprompt: Leveraging prompts to generate synthetic data for zero-shot relation triplet extraction, in: S. Muresan, P. Nakov, A. Villavicencio (Eds.), Findings of the Association for Computational Linguistics: ACL 2022, Dublin, Ireland, May 22-27, 2022, Association for Computational Linguistics, 2022, pp. 45–57. doi:10.18653/V1/2022.FINDINGS-ACL.5.
- [4] C. Chen, C. Li, ZS-BERT: towards zero-shot relation extraction with attribute representation learning, in: K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tür, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, Y. Zhou (Eds.), Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021, Association for Computational Linguistics, 2021, pp. 3470–3479. doi:10.18653/V1/2021.NAACL-MAIN.272.
- [5] D. Sorokin, I. Gurevych, Context-aware representations for knowledge base relation extraction, in: M. Palmer, R. Hwa, S. Riedel (Eds.), Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, EMNLP 2017, Copenhagen, Denmark, September 9-11, 2017, Association for Computational Linguistics, 2017, pp. 1784–1789. doi:10.18653/V1/D17-1188.
- [6] T. Rocktäschel, E. Grefenstette, K. M. Hermann, T. Kociský, P. Blunsom, Reasoning about entailment with neural attention, in: Y. Bengio, Y. LeCun (Eds.), 4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings, 2016. URL: <http://arxiv.org/abs/1509.06664>.
- [7] C. Chen, C. Li, ZS-BERT: towards zero-shot relation extraction with attribute representation learning, in: K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tür, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, Y. Zhou (Eds.), Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021, Association for Computational Linguistics, 2021, pp. 3470–3479. doi:10.18653/V1/2021.NAACL-MAIN.272.
- [8] V.-H. Tran, H. Ouchi, T. Watanabe, Y. Matsumoto, Improving discriminative learning for zero-shot relation extraction, in: Proceedings of the 1st Workshop on Semiparametric Methods in NLP: Decoupling Logic from Knowledge, 2022, pp. 1–6.
- [9] J. Zhao, W. Zhan, X. Zhao, Q. Zhang, T. Gui, Z. Wei, J. Wang, M. Peng, M. Sun, Re-matching: A fine-grained semantic matching method for zero-shot relation extraction, in: A. Rogers, J. L. Boyd-Graber, N. Okazaki (Eds.), Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), ACL 2023, Toronto, Canada, July 9-14, 2023, Association for Computational Linguistics, 2023, pp. 6680–6691. doi:10.18653/V1/2023.ACL-LONG.369.
- [10] B. Lv, X. Liu, S. Dai, N. Liu, F. Yang, P. Luo, Y. Yu, DSP: discriminative soft prompts for zero-shot entity and relation extraction, in: A. Rogers, J. L. Boyd-Graber, N. Okazaki (Eds.), Findings of the Association for Computational Linguistics: ACL 2023, Toronto, Canada, July 9-14, 2023, Association for Computational Linguistics, 2023, pp. 5491–5505. doi:10.18653/V1/2023.FINDINGS-ACL.339.
- [11] V.-H. Tran, H. Ouchi, H. Shindo, Y. Matsumoto, T. Watanabe, Enhancing semantic correlation between instances and relations for zero-shot relation extraction, *Journal of Natural Language Processing* 30 (2023) 304–329.
- [12] T. Ayoola, S. Tyagi, J. Fisher, C. Christodoulopoulos, A. Pierleoni, Refined: An efficient zero-

- shot-capable approach to end-to-end entity linking, in: A. Loukina, R. Gangadharaiah, B. Min (Eds.), Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: Industry Track, NAACL 2022, Hybrid: Seattle, Washington, USA + Online, July 10-15, 2022, Association for Computational Linguistics, 2022, pp. 209–220. doi:10.18653/V1/2022.NAAACL-INDUSTRY.24.
- [13] D. Zeng, K. Liu, S. Lai, G. Zhou, J. Zhao, Relation classification via convolutional deep neural network, in: J. Hajic, J. Tsujii (Eds.), COLING 2014, 25th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, August 23-29, 2014, Dublin, Ireland, ACL, 2014, pp. 2335–2344. URL: <https://aclanthology.org/C14-1220/>.
- [14] M. Miwa, M. Bansal, End-to-end relation extraction using lstms on sequences and tree structures, in: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, ACL 2016, August 7-12, 2016, Berlin, Germany, Volume 1: Long Papers, The Association for Computer Linguistics, 2016. doi:10.18653/V1/P16-1105.
- [15] Z. Zhong, D. Chen, A frustratingly easy approach for entity and relation extraction, in: K. Toutanova, A. Rumshisky, L. Zettlemoyer, D. Hakkani-Tür, I. Beltagy, S. Bethard, R. Cotterell, T. Chakraborty, Y. Zhou (Eds.), Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2021, Online, June 6-11, 2021, Association for Computational Linguistics, 2021, pp. 50–61. doi:10.18653/V1/2021.NAAACL-MAIN.5.
- [16] S. Wu, Y. He, Enriching pre-trained language model with entity information for relation classification, in: W. Zhu, D. Tao, X. Cheng, P. Cui, E. A. Rundensteiner, D. Carmel, Q. He, J. X. Yu (Eds.), Proceedings of the 28th ACM International Conference on Information and Knowledge Management, CIKM 2019, Beijing, China, November 3-7, 2019, ACM, 2019, pp. 2361–2364. doi:10.1145/3357384.3358119.
- [17] J. Ni, G. Rossiello, A. Gliozzo, R. Florian, A generative model for relation extraction and classification, CoRR abs/2202.13229 (2022). URL: <https://arxiv.org/abs/2202.13229>. arXiv:2202.13229.
- [18] X. Chen, N. Zhang, X. Xie, S. Deng, Y. Yao, C. Tan, F. Huang, L. Si, H. Chen, Knowprompt: Knowledge-aware prompt-tuning with synergistic optimization for relation extraction, in: F. Laforest, R. Troncy, E. Simperl, D. Agarwal, A. Gionis, I. Herman, L. Médini (Eds.), WWW '22: The ACM Web Conference 2022, Virtual Event, Lyon, France, April 25 - 29, 2022, ACM, 2022, pp. 2778–2788. doi:10.1145/3485447.3511998.
- [19] C. Wu, L. Chen, Utber: Utilizing fine-grained entity types to relation extraction with distant supervision, in: IEEE International Conference on Smart Data Services, SMDS 2020, Beijing, China, October 19-23, 2020, IEEE, 2020, pp. 63–71. doi:10.1109/SMDS49396.2020.00015.
- [20] M. Koch, J. Gilmer, S. Soderland, D. S. Weld, Type-aware distantly supervised relation extraction with linked arguments, in: A. Moschitti, B. Pang, W. Daelemans (Eds.), Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL, ACL, 2014, pp. 1891–1901. doi:10.3115/V1/D14-1203.
- [21] Y. Liu, K. Liu, L. Xu, J. Zhao, Exploring fine-grained entity type constraints for distantly supervised relation extraction, in: J. Hajic, J. Tsujii (Eds.), COLING 2014, 25th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, August 23-29, 2014, Dublin, Ireland, ACL, 2014, pp. 2107–2116. URL: <https://aclanthology.org/C14-1199/>.
- [22] M. Rahimi, M. Surdeanu, Improving zero-shot relation classification via automatically-acquired entailment templates, in: B. Can, M. Mozes, S. Cahyawijaya, N. Saphra, N. Kassner, S. Ravfogel, A. Ravichander, C. Zhao, I. Augenstein, A. Rogers, K. Cho, E. Grefenstette, L. Voita (Eds.), Proceedings of the 8th Workshop on Representation Learning for NLP, Repl4NLP@ACL 2023, Toronto, Canada, July 13, 2023, Association for Computational Linguistics, 2023, pp. 187–195. doi:10.18653/V1/2023.REPL4NLP-1.16.
- [23] O. Sainz, O. L. de Lacalle, G. Labaka, A. Barrena, E. Agirre, Label verbalization and entailment for effective zero and few-shot relation extraction, in: M. Moens, X. Huang, L. Specia, S. W. Yih (Eds.), Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing,

- EMNLP 2021, Virtual Event / Punta Cana, Dominican Republic, 7-11 November, 2021, Association for Computational Linguistics, 2021, pp. 1199–1212. doi:10.18653/v1/2021.EMNLP-MAIN.92.
- [24] A. Obamuyide, A. Vlachos, Zero-shot relation classification as textual entailment, in: Proceedings of the first workshop on fact extraction and VERification (FEVER), 2018, pp. 72–78.