Few-Shot Event Argument Extraction Based on a Meta-Learning Approach

Aboubacar Tuo, Romaric Besançon, Olivier Ferret, Julien Tourille
Université Paris-Saclay, CEA, List, F-91120, Palaiseau, France
{aboubacar.tuo,romaric.besancon,olivier.ferret,julien.tourille}@cea.fr

Abstract

Few-shot learning techniques for Event Extraction are developed to alleviate the cost of data annotation. However, most studies on few-shot event extraction only focus on event trigger detection and no study has been proposed on argument extraction in a meta-learning context. In this paper, we investigate few-shot event argument extraction using prototypical networks, casting the task as a relation classification problem and adapting the classical N-ways K-shots approach in such a way that new classes correspond to new types of events, which is more realistic than focusing on types of arguments. Furthermore, we propose to enhance the relation embeddings by injecting syntactic knowledge into the model using graph convolutional networks. Our experimental results, obtained in an evaluation framework specifically designed for meta-learning approaches, show that our proposed approach achieves strong performance on ACE 2005 in several few-shot configurations and highlight the importance of syntactic knowledge for this task.

1 Introduction

Event Extraction (EE) aims to automatically identify and extract information about events from unstructured texts, targeting more specifically the event trigger (the word or phrase corresponding to the mention of the event) and the event arguments (the entities that play a role in the event). For instance, in the sentence "*Seven U.S. soldiers were killed when their vehicle hit an explosive device in Baghdad*", a *Life.Die* event, according to the ACE 2005 (Walker et al., 2006) terminology for event types, is mentioned through the trigger *killed* and associated with the arguments *Seven U.S. soldiers*, *explosive device*, and *Baghdad*, which correspond to the respective roles of victim, instrument, and location of the event structure.

Typical EE systems rely on supervised approaches that require a large amount of annotated data for each considered event type. Unfortunately, data annotation is expensive and cannot be performed for all applications, since new event types may appear with only a few examples. As a result, there has been a growing interest in addressing the challenge of Few-Shot Event Extraction (FSEE).

Most studies in FSEE only consider event detection (ED), which focuses on extracting and classifying event triggers. Some of them further restrict the ED task to the classification part only, using a candidate trigger already identified (Lai et al., 2021, 2020b,a; Deng et al., 2020). Other leverage event-related keywords (Bronstein et al., 2015; Lai and Nguyen, 2019) or external resources for data enrichment (Deng et al., 2021; Zhang et al., 2021; Shen et al., 2021). Prototypical networks (Snell et al., 2017) have also been applied successfully to this task formalized as a sequence annotation problem in a meta-learning context (Cong et al., 2021; Tuo et al., 2022, 2023).

However, very few studies address argument extraction using these methods. Most existing methods for event argument extraction in low-resource scenarios rely on supervised approaches, with additional experiments to show the performance of the models with limited annotated data. Such approaches exploit question answering frameworks (Du and Cardie, 2020; Zhou et al., 2021), specific slot-filling techniques (Chen et al., 2020; Hsu et al., 2022; Dai et al., 2022; Ma et al., 2022), or methods based on textual entailment (Sainz et al., 2022). Zero-shot approaches have also been proposed, either relying on external resources (Huang et al., 2018; Zhang et al., 2021) or using prompting techniques with pre-trained language models (Lin et al., 2023). To the best of our knowledge, only Yang et al. (2023) tackle few-shot argument extraction, but they do it at the document level and the performance of their approach is rather limited.

Our work focuses on using prototypical networks for FSEE, modeling the task as a relation

Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 4: Student Research Workshop), pages 146–153
June 18, 2024 ©2024 Association for Computational Linguistics

classification problem between candidate entities and the event trigger. Moreover, we propose several extensions of this model by injecting syntactic information into the representation of relations. Our contributions are summarized as follows:

- we devise a new approach for few-shot event argument extraction that yields encouraging results and offers a new evaluation framework for few-shot event argument extraction at the sentence level;
- we cast the few-shot event argument extraction task as a relation classification problem;
- we highlight the benefits of integrating syntactic knowledge.

2 Proposed Method

We tackle event argument extraction as a relation classification task between an event trigger and the entities in the sentence containing the trigger.[1] More precisely, for each of these entities, we perform a multi-class classification, each class corresponding to one of the possible roles for the event type associated with the considered trigger. In addition, a null class stands for the absence of a role for a candidate entity with respect to the event.

Our few-shot setting for this task is a variant of the standard N-ways K-shots meta-learning approach (Vinyals et al., 2016), which involves learning through multiple training episodes. Each episode represents a classification task $\mathcal{T} = \{\mathcal{S}, \mathcal{Q}\}$, with a support set $\mathcal{S}$ and a query set $\mathcal{Q}$, where $\mathcal{S}$ contains N classes, each with K labeled instances, and $\mathcal{Q}$ contains examples from the same N classes as $\mathcal{S}$. The goal of an episode is to classify the query set examples based on the support set examples. The idea is to train the model on various episodes of different tasks so that it can quickly adapt and generalize to new tasks (during inference), including on previously unseen argument role classes.

For a given event type e, an instance is designed as $x_i = (s_i^e, tr_i^e, e_i)$, with $y_i = a_i$ its label, where s_i^e is the sentence mentioning the event, tr_i^e the trigger, e_i the argument candidate, and a_i the role belonging to $\mathcal{A}^e = \{\mathcal{A}_+^e \bigcup None\}$ with $\mathcal{A}_+^e$ the argument set of the event type e and $None$ denotes that the entity has no role in the event. The same sentence can therefore belong to as many instances as it contains entities.

[1] In this study, we restrict ourselves to the detection of event arguments within the same sentence.

Our formulation differs slightly from the typical meta-learning setting as, even if the classification is performed on the event argument roles, we consider new classes at the event level (the event types in the test set were not encountered during training). We therefore have new event types during inference, but we may have seen argument roles with the same semantics or not. This is more in line with real-world applications where new events can arise instead of just new roles for existing events. As a consequence, our formulation of the N-ways K-shot setup includes a variable N, which represents the number of arguments for a given event type. Also, some arguments in the events of the test set may be similar to arguments from the training set, even if they correspond to different event types (e.g. an argument *Agent* may exist in several events).

2.1 Overall Framework

An overview of our approach is given in Figure 1. Our model takes as input an instance $x_i = (s_i^e, tr_i^e, e_i)$, which is processed by an encoder to produce a hidden representation h_i. This representation is then classified using a metric-based meta-learning algorithm.

2.1.1 Instance Encoder

The encoder builds an embedding from an instance $h_i = \mathcal{E}(x_i)$. To help the encoder focus on the trigger and the entity within a sentence, we mark them with special tokens, as proposed in previous works (Zhang et al., 2019; Han et al., 2018; Baldini Soares et al., 2019). Then, we feed the whole sentence to an encoder-like language model to obtain an embedding of each word, using **BERT** (Devlin et al., 2019) as a baseline. Finally, we concatenate the embeddings of the trigger and entity heads to obtain h_i, our embedding of the role.

Syntax integration. We also propose to improve the role embedding by leveraging syntactic information, since the relation between the event trigger and the event argument in the sentence is often expressed through a syntactic relation. Moreover, integrating syntactic information previously showed improved performance for event extraction (Nguyen and Grishman, 2018; Balali et al., 2020), either when arguments are distant from their trigger or when roles can be ambiguous. For instance, in the context of a *Transport* event, the *origin* and *destination* locations can be easily confused. More precisely, we propose two ways to integrate syntactic information into role embeddings.

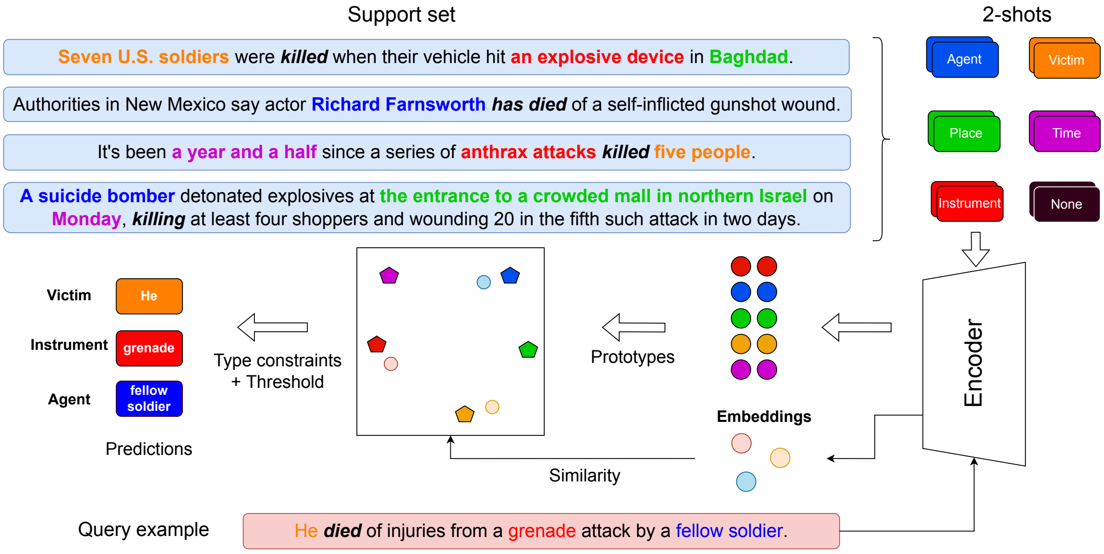

Figure 1: Overview of our method. Trigger words are in ***bold italic***.

Encoder		5 shots P	5 shots R	5 shots F1	10 shots P	10 shots R	10 shots F1
BERT	Proto	63.1 ± 0.9	56.4 ± 1.0	59.6 ± 0.5	66.4 ± 0.5	61.6 ± 0.7	63.9 ± 0.3
	C-Proto	62.7 ± 0.9	57.0 ± 1.2	60.0 ± 1.0	67.1 ± 0.8	63.8 ± 0.9	65.5 ± 0.8
BERT++	Proto	64.9 ± 1.1	58.6 ± 1.2	61.6 ± 0.8	66.8 ± 1.5	63.8 ± 1.1	65.2 ± 0.6
	C-Proto	65.8 ± 0.5	58.8 ± 1.8	62.1 ± 1.0	66.8 ± 1.7	**66.5*** ± **1.7**	**66.7*** ± **1.0**
RGCN	Proto	**69.0** ± **2.1**	56.6 ± 4.0	62.2 ± 2.2	**71.2*** ± **0.7**	60.0 ± 1.5	65.0 ± 0.9
	C-Proto	68.5 ± 1.1	**59.2*** ± **1.7**	**63.5*** ± **1.2**	69.2 ± 0.5	61.4 ± 0.8	65.1 ± 0.5

Table 1: Event argument extraction results: Precision (P), Recall (R), and F1-score (F1). Our best scores are in **bold** and the second best are underlined. * when the best score is statistically higher than the second one.

BERT++ performs this integration in a static manner via a look-up matrix: embeddings of the entity's Part-of-Speech (PoS) tags and the syntactic dependency path between the trigger and the entity are concatenated to the trigger and entity embeddings. We address the variable-length nature of dependency paths by applying max-pooling to the representations of their dependencies.

RGCN exploits a dynamic integration of syntactic information using Relational Graph Convolution Networks (Schlichtkrull et al., 2018) on the dependency path: the sentence is passed to the BERT encoder followed by L layers of RGCN and the instance representation is the concatenation of the resulting embeddings for the trigger and candidate entity.

2.1.2 Classification Module

The classification module aims to classify instances based on their similarity to each class representation. In this framework, the classification of an instance is performed by comparing its representation with the class prototypes. We performed experiments with Prototypical Networks and Contrastive Prototypical Networks.

Prototypical Networks (Proto) are based on the idea of learning a prototype representation for each class in the training set. During training, the encoder is used to convert the instances into embeddings and the prototype of a class is simply defined as an aggregation of the embeddings of its associated instances (generally the mean). During the test phase, the predicted class for an input instance is chosen as its nearest prototype.

Contrastive Prototypical Networks (C-Proto) offer an alternative to the standard prototypical networks, designed to solve the problem of the null class. Since all entities in the sentences are classified, the entities that do not hold any role in the event must be associated with the null class, which does not have any real semantics. Following previous work (Tan et al., 2019; Tuo et al., 2023), we use a contrastive loss function to train the model

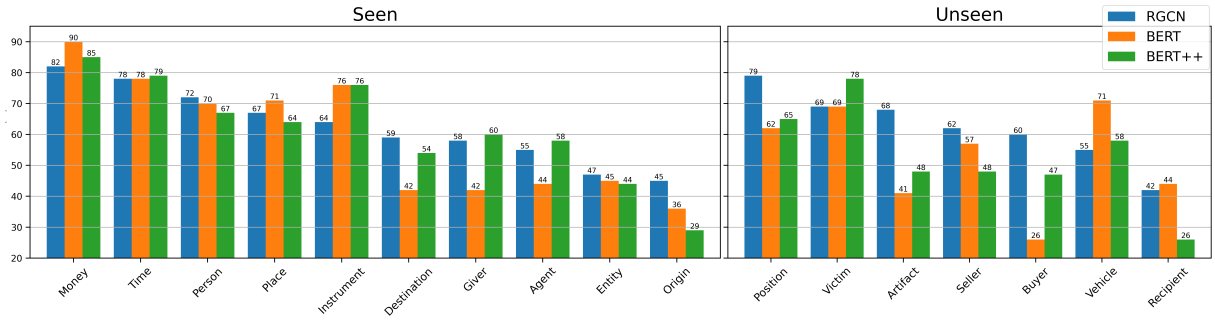

Figure 2: F1-score on event argument extraction by role.

to separate the null examples from the examples of true classes. Then, we find a decision threshold to decide whether an example is part of an argument class or not. Unlike Tuo et al. (2023), who use a cumulative density function, we compute the threshold using the similarity value found on the closest example in the support set.

Similarly to previous work (Sainz et al., 2022; Lin et al., 2023), we also use the entity type knowledge to constrain the predicted arguments, since this information is available in the considered dataset (the annotation guidelines of ACE 2005 contain the possible entity types for each event role). We integrate this domain knowledge by selecting the nearest class that has a compatible entity type for the role.

3 Experiments

Settings. We conduct our experiments on the ACE-2005 dataset with the split provided by Lai et al. (2021). This split ensures that there is no overlap between train and test classes, thus simulating a real-life few-shot scenario.

We use BERT-large-uncased as our backbone encoder. Additionally, for BERT++ encoder, we use trainable vectors of size 256 to encode syntactic dependencies and PoS tags obtained using spaCy. The RGCN encoder is composed of two convolutional layers and three relation types (i.e. direct paths, indirect paths, and self-loops). We provide a list of hyper-parameters in Table 2.

Main Results. Our main results are presented in Table 1. The entities considered for argument extraction are the gold entities from ACE 2005. Globally, integrating syntactic knowledge improves the performance in all cases and richer syntactic information using RGCN is better when little data is available (5 shots). These results also show a gain in performance when using contrastive learning compared with the vanilla Prototypical Networks, which confirms previous results on ED (Tuo et al., 2023), but to a lesser extent.

Parameter	**Value**
base encoder	BERT-large-uncased
sequence length	128
train iteration	5,000
optimizer	AdamW
learning rate	
BERT	$1e-5$
Others	$1e-4$
Weight decay	$1e-2$
dropout	0.1
warmup ratio	0.1
scheduler	StepLR
β_1 β_2	0.9 0.999
RGCN layers	2

Table 2: Hyperparameters.

Detailed results for each event role are presented in Figure 2 and show that the RGCN model mainly contributes to roles that can be mixed up, such as the *Origin* and *Destination* roles in a *Transport* event. In contrast, it hurts less ambiguous roles, such as *Instrument* or *Vehicle*. Indeed, syntactic information is particularly useful to disambiguate similar/symmetric roles in the same event, whereas our baseline model based on the simple concatenation of the trigger and entity embeddings is not sufficient to dispel the confusion. Figure 2 also shows that the interest of syntactic information is observed both for roles seen during training and new roles.

Figure 3 provides an overview of the representations learned by each encoder on the evaluation set. The pre-trained BERT encoder corresponds to a BERT model without any fine-tuning specific to the event argument extraction task. We compare the three encoders presented in this work: BERT, BERT++, and RGCN. First, it is clear that training the BERT model significantly improves the

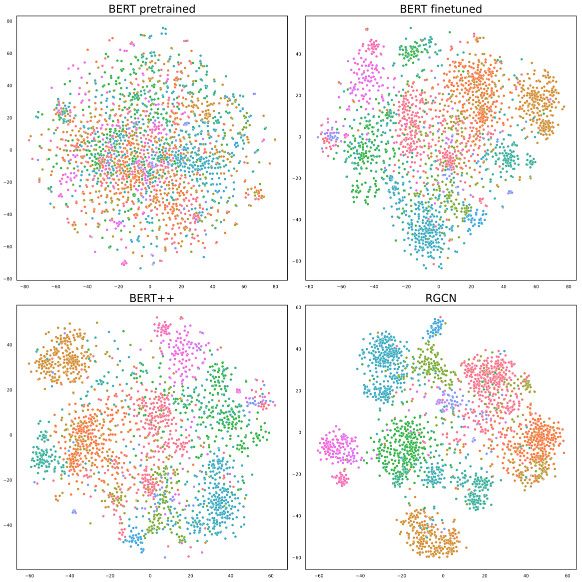

Figure 3: Visualization of argument embeddings using t-SNE. Each point represents an argument and its color corresponds to its role.

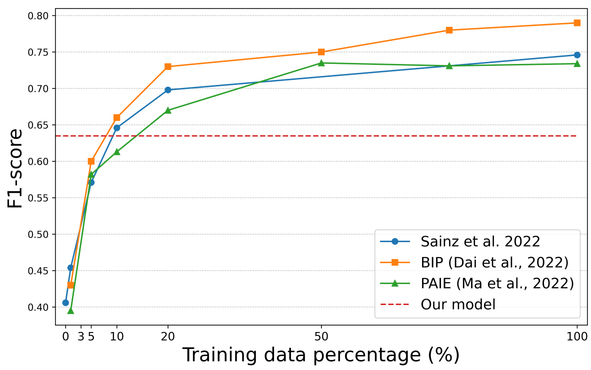

Figure 4: Comparison of our best model, RGCN C-Proto (dot line), in the 5-shots configuration to state-of-the-art models with few annotated data.

possibility to discriminate among argument types compared to an untuned BERT. This highlights the relevance of the model we have adopted for this task and the importance of fine-tuning the BERT model in this context. We can also observe qualitatively that the two syntactically enriched encoders, BERT++ and RGCN, seem to provide more discriminative representations than the original BERT encoder. These observations are consistent with the results obtained during the evaluation, which reinforces the relevance of enriching the representations with syntactic information and suggests an overall improvement in the model's performance.

Comparison to the state-of-the-art. Since we propose a new framework for few-shot event argument extraction using meta-learning episodic evaluation, the results are not directly comparable with previous studies, which use other configurations. However, we propose a comparison with works that make similar assumptions about the available input knowledge and require a comparable amount of annotated data.

More precisely, we compare our approach to PAIE (Ma et al., 2022), BIP (Dai et al., 2022), and NLI (Sainz et al., 2022), which also assign roles to gold entities, but with a different formulation of "few-shot". Our approach lies in defining new classes with a limited amount of data while theirs consider learning with a small amount of data, but cannot be evaluated on new unseen classes. We consider the class transfer approach to be more realistic in a real-world context, as few-shot requirements are driven precisely by the emergence of new event types.

To perform a form of comparison, we focus on the amount of annotated data for the event types in the ACE 2005 test set. Figure 4 shows the evolution of the performance for event argument extraction for three baseline models we consider as a function of the percentage of training data. Together with these curves, we have plotted the performance level of our 5-shots configuration (i.e., 5 examples per role), which corresponds in quantity to about 3% of the evaluation data. However, it should be noted that our model is trained on 18 event types and all their examples, the 3% of data only concerning the types not seen during training. Nevertheless, Figure 4 shows that up to 5% of the training data, our proposed model outperforms the considered baseline models. We can therefore conclude that our meta-learning approach is particularly well suited to a regime in which very little annotated data is available for the target event types.

Ablation study. We present in Table 3 an ablation study for our best encoder, RGCN, with the C-Proto model, to investigate the respective impacts of using a dynamic threshold[2] and using entity type constraints on the predicted event roles. Table 3 shows that both the threshold and the constraints significantly contribute to the performance of the model, but differently: while the constraints favor recall, the threshold improves precision, the threshold having globally a higher impact than the constraints.

[2]To remove the use of the threshold, we rebuild a prototype for the null class during the test phase.

	P	R	F1
Full model	68.5	59.2	63.5
w/o threshold	47.9	59.3	52.9
w/o constraints	68.2	50.9	58.3
w/o threshold & constraints	33.9	61.9	43.8

Table 3: Ablation study in 5-shots setting.

Limitations

Since our approach requires entity information, it is not applicable in scenarios where entities are not provided. However, it may still be possible to adapt it for scenarios where entities are not explicitly provided. One potential solution is to incorporate an entity candidate extraction method upstream of our approach. These candidate entities can then be used as input to our method. However, this approach may introduce additional noise and errors due to the imperfect nature of named entity recognition methods.

Furthermore, our formulation only considers interactions between triggers in entities, leaving out entity-entity interactions, which are often valuable for event argument extraction (Sha et al., 2018). We leave this as a direction for future work.

4 Conclusion and Future Work

In this article, we propose a meta-learning approach for event argument extraction that complements existing work on few-shot event extraction. We cast the argument extraction as a relation classification problem and propose an adaptation of the N-ways K-shots framework that matches the expectations of a real-life few-shot event extraction task. We show that our proposed models achieve strong performance on the ACE 2005 dataset. Our experiments prove the interest of enhancing the event role embeddings using syntactic information.

As a perspective, we plan to extend the development of few-shot models for event extraction towards the definition of joint approaches integrating entity, trigger, and argument extractions more tightly.

Acknowledgments

This publication was made possible by the use of the FactoryIA supercomputer, financially supported by the Ile-de-France Regional Council.

References

Ali Balali, Masoud Asadpour, Ricardo Campos, and Adam Jatowt. 2020. Joint event extraction along shortest dependency paths using graph convolutional networks. *Knowledge-Based Systems*, 210:106492.

Livio Baldini Soares, Nicholas FitzGerald, Jeffrey Ling, and Tom Kwiatkowski. 2019. Matching the blanks: Distributional similarity for relation learning. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 2895–2905, Florence, Italy. Association for Computational Linguistics.

Ofer Bronstein, Ido Dagan, Qi Li, Heng Ji, and Anette Frank. 2015. Seed-based event trigger labeling: How far can event descriptions get us? In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 372–376, Beijing, China. Association for Computational Linguistics.

Yunmo Chen, Tongfei Chen, Seth Ebner, Aaron Steven White, and Benjamin Van Durme. 2020. Reading the manual: Event extraction as definition comprehension. In *Proceedings of the Fourth Workshop on Structured Prediction for NLP*, pages 74–83, Online. Association for Computational Linguistics.

Xin Cong, Shiyao Cui, Bowen Yu, Tingwen Liu, Wang Yubin, and Bin Wang. 2021. Few-Shot Event Detection with Prototypical Amortized Conditional Random Field. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 28–40, Online. Association for Computational Linguistics.

Lu Dai, Bang Wang, Wei Xiang, and Yijun Mo. 2022. Bi-directional iterative prompt-tuning for event argument extraction. In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, pages 6251–6263, Abu Dhabi, United Arab Emirates. Association for Computational Linguistics.

Shumin Deng, Ningyu Zhang, Jiaojian Kang, Yichi Zhang, Wei Zhang, and Huajun Chen. 2020. Meta-Learning with Dynamic-Memory-Based Prototypical Network for Few-Shot Event Detection. In *WSDM*, pages 151–159, Houston, TX, USA.

Shumin Deng, Ningyu Zhang, Luoqiu Li, Chen Hui, Tou Huaixiao, Mosha Chen, Fei Huang, and Huajun Chen. 2021. OntoED: Low-resource event detection with ontology embedding. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*, pages 2828–2839, Online. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of deep bidirectional transformers for language understanding. 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)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Xinya Du and Claire Cardie. 2020. Event extraction by answering (almost) natural questions. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 671–683, Online. Association for Computational Linguistics.

Xu Han, Hao Zhu, Pengfei Yu, Ziyun Wang, Yuan Yao, Zhiyuan Liu, and Maosong Sun. 2018. FewRel: A large-scale supervised few-shot relation classification dataset with state-of-the-art evaluation. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4803–4809, Brussels, Belgium. Association for Computational Linguistics.

I-Hung Hsu, Kuan-Hao Huang, Elizabeth Boschee, Scott Miller, Prem Natarajan, Kai-Wei Chang, and Nanyun Peng. 2022. DEGREE: A data-efficient generation-based event extraction model. In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1890–1908, Seattle, United States. Association for Computational Linguistics.

Lifu Huang, Heng Ji, Kyunghyun Cho, Ido Dagan, Sebastian Riedel, and Clare Voss. 2018. Zero-shot transfer learning for event extraction. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 2160–2170, Melbourne, Australia. Association for Computational Linguistics.

Viet Lai, Franck Dernoncourt, and Thien Huu Nguyen. 2021. Learning prototype representations across few-shot tasks for event detection. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 5270–5277, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

Viet Dac Lai, Franck Dernoncourt, and Thien Huu Nguyen. 2020a. Exploiting the Matching Information in the Support Set for Few Shot Event Classification. In Hady W. Lauw, Raymond Chi-Wing Wong, Alexandros Ntoulas, Ee-Peng Lim, See-Kiong Ng, and Sinno Jialin Pan, editors, *Advances in Knowledge Discovery and Data Mining*, volume 12085, pages 233–245. Springer, Cham. Series Title: Lecture Notes in Computer Science.

Viet Dac Lai and Thien Nguyen. 2019. Extending event detection to new types with learning from keywords. In *Proceedings of the 5th Workshop on Noisy User-generated Text (W-NUT 2019)*, pages 243–248, Hong Kong, China. Association for Computational Linguistics.

Viet Dac Lai, Thien Huu Nguyen, and Franck Dernoncourt. 2020b. Extensively matching for few-shot learning event detection. In *Proceedings of the First Joint Workshop on Narrative Understanding, Storylines, and Events*, pages 38–45, Online. Association for Computational Linguistics.

Zizheng Lin, Hongming Zhang, and Yangqiu Song. 2023. Global constraints with prompting for zero-shot event argument classification. In *Findings of the Association for Computational Linguistics: EACL 2023*, pages 2527–2538, Dubrovnik, Croatia. Association for Computational Linguistics.

Yubo Ma, Zehao Wang, Yixin Cao, Mukai Li, Meiqi Chen, Kun Wang, and Jing Shao. 2022. Prompt for extraction? PAIE: Prompting argument interaction for event argument extraction. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 6759–6774, Dublin, Ireland. Association for Computational Linguistics.

Thien Nguyen and Ralph Grishman. 2018. Graph convolutional networks with argument-aware pooling for event detection. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1).

Oscar Sainz, Itziar Gonzalez-Dios, Oier Lopez de Lacalle, Bonan Min, and Eneko Agirre. 2022. Textual entailment for event argument extraction: Zero- and few-shot with multi-source learning. In *Findings of the Association for Computational Linguistics: NAACL 2022*, pages 2439–2455, Seattle, United States. Association for Computational Linguistics.

Michael Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, and Max Welling. 2018. Modeling relational data with graph convolutional networks. In *The Semantic Web*, pages 593–607, Cham. Springer International Publishing.

Lei Sha, Feng Qian, Baobao Chang, and Zhifang Sui. 2018. Jointly extracting event triggers and arguments by dependency-bridge rnn and tensor-based argument interaction. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1).

Shirong Shen, Tongtong Wu, Guilin Qi, Yuan-Fang Li, Gholamreza Haffari, and Sheng Bi. 2021. Adaptive knowledge-enhanced Bayesian meta-learning for few-shot event detection. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 2417–2429, Online. Association for Computational Linguistics.

Jake Snell, Kevin Swersky, and Richard Zemel. 2017. Prototypical networks for few-shot learning. In *Advances in Neural Information Processing Systems*, volume 30.

Ming Tan, Yang Yu, Haoyu Wang, Dakuo Wang, Saloni Potdar, Shiyu Chang, and Mo Yu. 2019. Out-of-domain detection for low-resource text classification tasks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing* *and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3566–3572, Hong Kong, China. Association for Computational Linguistics.

Aboubacar Tuo, Romaric Besançon, Olivier Ferret, and Julien Tourille. 2022. Better exploiting bert for few-shot event detection. In *Natural Language Processing and Information Systems*, pages 291–298, Cham. Springer International Publishing.

Aboubacar Tuo, Romaric Besançon, Olivier Ferret, and Julien Tourille. 2023. Trigger or not trigger: Dynamic thresholding for few shot event detection. In *Advances in Information Retrieval*, pages 637–645, Cham. Springer Nature Switzerland.

Oriol Vinyals, Charles Blundell, Timothy Lillicrap, koray kavukcuoglu koray, and Daan Wierstra. 2016. Matching networks for one shot learning. In *Advances in Neural Information Processing Systems*, volume 29.

Christopher Walker, Stephanie Strassel, and Kazuaki Maeda Julie Medero. 2006. ACE 2005 Multilingual Training Corpus.

Xianjun Yang, Yujie Lu, and Linda Petzold. 2023. Few-shot document-level event argument extraction. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 8029–8046, Toronto, Canada. Association for Computational Linguistics.

Hongming Zhang, Haoyu Wang, and Dan Roth. 2021. Zero-shot Label-aware Event Trigger and Argument Classification. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, pages 1331–1340, Online. Association for Computational Linguistics.

Zhengyan Zhang, Xu Han, Zhiyuan Liu, Xin Jiang, Maosong Sun, and Qun Liu. 2019. ERNIE: Enhanced language representation with informative entities. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 1441–1451, Florence, Italy. Association for Computational Linguistics.

Yang Zhou, Yubo Chen, Jun Zhao, Yin Wu, Jiexin Xu, and Jinlong Li. 2021. What the role is vs. what plays the role: Semi-supervised event argument extraction via dual question answering. *Proceedings of the AAAI Conference on Artificial Intelligence*, 35(16):14638–14646.