

Reinforcement-based denoising of distantly supervised NER with partial annotation

Farhad Nooralahzadeh, Jan Tore Lønning, Lilja Øvrelid

Department of Informatics

University of Oslo, Norway

{farhadno, jtl, liljao}@ifi.uio.no

Abstract

Existing named entity recognition (NER) systems rely on large amounts of human-labeled data for supervision. However, obtaining large-scale annotated data is challenging particularly in specific domains like health-care, e-commerce and so on. Given the availability of domain specific knowledge resources, (e.g., ontologies, dictionaries), distant supervision is a solution to generate automatically labeled training data to reduce human effort. The outcome of distant supervision for NER, however, is often noisy. False positive and false negative instances are the main issues that reduce performance on this kind of auto-generated data. In this paper, we explore distant supervision in a supervised setup. We adopt a technique of partial annotation to address false negative cases and implement a reinforcement learning strategy with a neural network policy to identify false positive instances. Our results establish a new state-of-the-art on four benchmark datasets taken from different domains and different languages. We then go on to show that our model reduces the amount of manually annotated data required to perform NER in a new domain.

1 Introduction

Named Entity Recognition (NER) is one of the primary tasks in information extraction pipelines. (Ma and Hovy, 2016; Lample et al., 2016; Peters et al., 2018; Akbik et al., 2018). Traditional studies apply statistical techniques such as Hidden Markov Models (HMM) and Conditional Random Fields (CRF) using large amounts of features and extra resources (Ratinov and Roth, 2009; Passos et al., 2014). In recent years, deep learning approaches achieve state-of-the-art results in the task without any feature engineering (Ma and Hovy, 2016; Lample et al., 2016). Most of these works assume that there is a certain amount of annotated sentences in the training phase. However, avail-

ability of large amounts of labeled data is problematic, particularly in specific domains. Distant supervision is proposed by Mintz et al. (2009) to address the challenge of obtaining training data for new domains using existing knowledge resources (dictionaries, ontologies). It has previously been successfully applied to tasks like relation extraction (Riedel et al., 2010; Augenstein et al., 2014) and entity recognition (Fries et al., 2017; Shang et al., 2018b; Yang et al., 2018). For the task of NER, it identifies entity mentions if it exist in the knowledge base (e.g. domain-specific dictionary, glossary, ontology) and assigns the corresponding type according to the knowledge base.

However, distant supervision approaches encounter two main limitations. First, due to limited coverage of the knowledge resources, unmatched tokens result in False Negatives (FNs). Second, since simple string matching is employed to detect entity mentions, ambiguity in the knowledge resource may lead to False Positives (FPs). For the FN problem, Tsuboi et al. (2008) incorporate partial annotations into CRFs and propose a parameter estimation method for CRFs using partially annotated corpora (here-in after referred to as Partial-CRF). In order to reduce the negative impact of FPs for relation extraction, Qin et al. (2018) propose a deep reinforcement learning (RL) agent where the the agent's goal is to decide whether to remove or keep the distantly supervised instance.

In this paper we make the following contributions:

- We combine the Partial-CRF approach with performance-driven, policy-based reinforcement learning to clean the noisy, distantly supervised data for NER in a pre-processing step.
- We formulate the reward function in RL based on the change in the performance of

the NER module where the policy of RL is trained in an unsupervised manner by interaction with the environment.

- We show that our approach can boost the performance of the neural NER system on four datasets from different domains and for two different languages (English and Chinese).

2 Related work

The task of NER has been widely studied in the last decade and is generally considered as a sequence labeling problem. Using neural techniques, many studies report state-of-the-art results on this type of sequence labeling task (Lample et al., 2016; Ma and Hovy, 2016). These types of studies utilize character and/or word embeddings to encode sentence-level features automatically. Recently, the use of contextualized word representation (Peters et al., 2018; Akbik et al., 2018) significantly improves the state-of-the-art results in many sequence labeling tasks and specifically also in the NER benchmark.

In the supervised NER paradigm, this task suffers from lack of large-scale labeled training data when moving to a new domain or new language. To alleviate the reliance on human annotated data, distant supervision is proposed by Mintz et al. (2009), to generate annotated data by heuristically aligning text to an existing domain-specific knowledge resource. It is widely used for relation extraction (Mintz et al., 2009; Riedel et al., 2010; Augenstein et al., 2014) and lately it has attracted attention also for NER (Ren et al., 2015; Fries et al., 2017; Shang et al., 2018b; Yang et al., 2018). Shang et al. (2018b) present the AutoNER model which employs a new type of tagging scheme (i.e., Tie or Break) rather than common ones (i.e., IOB, IOBES) without any CRF layer and achieves state-of-the-art unsupervised *F1* scores on several benchmark datasets. Crucially, they employ a set of high-quality phrases in distant supervision, using a phrase mining technique (Shang et al., 2018a) to reduce the false-negative labels. Feng et al. (2018) and Yang et al. (2018) make use of reinforcement learning to tackle false positives in distantly supervised relation classification and NER, respectively. Similar to our work, Yang et al. (2018) address the noisy automatic annotation in NER, by using partial annotation learning and reinforcement learning. However, unlike our approach, they train

the NER model and reinforcement learning model jointly, calculating the reward based on the loss of the NER model, whereas we employ the RL module as a pre-processing/filtering step, incorporating the previous state to satisfy a Markov decision process (MDP). Yang et al. (2018) evaluate only on a Chinese dataset, whereas we apply our model also to English datasets. Furthermore, after running their code¹, we observe that to reach the reported results in their paper on e-commerce dataset, the model needs more than 500 epochs and the reinforcement learning component removes all the distantly annotated sentences after some epochs. It means that after some epochs the code performs only the base-line NER model on annotation dataset and ignoring RL module, since there are no distantly annotated sentences. Their two datasets are included in our experiment in order to compare to their results. Qin et al. (2018) explore deep reinforcement learning as a false positive removal tool for distantly supervised relation extraction. Here, we adapt their approach to the NER task. Unlike Qin et al. (2018) however, we learn the policy agent in an unsupervised manner, where the parameters are learnt by interaction with the environment.

3 Model

We implement Partial-CRF together with a performance-driven, policy-based reinforcement learning method to detect FNs and FPs in distantly supervised NER. In contrast to a previous study that has applied RL in NER (Yang et al., 2018), we consider the RL agent as a pre-processing task to clean FPs from the noisy dataset. Furthermore, our RL agent is rewarded based on the change in the performance of the NER module and it is modeled as a Markov decision process (MDP).

Algorithm 1 describes the overall training procedure for our model and in the following, we detail the various components of our model.

3.1 Baseline NER model

The goal of NER is to identify text spans that present named entities and assign them into pre-defined categories. These categories vary depending on the domain, for example in the general domain, they are categories like organization, person and location names; in bio-medical domain,

¹<https://github.com/rainarch/DSNER>

B-Disease	B-Disease	B-Disease	B-Disease	B-Disease									
I-Disease	I-Disease	I-Disease	I-Disease	I-Disease									
E-Disease	E-Disease	E-Disease	E-Disease	E-Disease									
S-Disease	S-Disease	S-Disease	S-Disease	S-Disease									
B-Chemical	B-Chemical	B-Chemical	B-Chemical	B-Chemical									
I-Chemical	I-Chemical	I-Chemical	I-Chemical	I-Chemical									
E-Chemical	E-Chemical	E-Chemical	E-Chemical	E-Chemical									
S-Chemical	S-Chemical	S-Chemical	S-Chemical	S-Chemical									
O	O	O	O	O	B-Disease	E-Disease	O	O	O	O	O	O	S-Chemical
leprosy	developed	a	Heinz	body	hemolytic	anemia	while	taking	a	dose	of	dapsona	

Figure 1: Annotation of distantly labeled example in Partial CRF based on IOBES scheme. The words with green tags are found in dictionary and assigned to the corresponding entity types, and the ones that are not found in dictionary are assigned to all possible tags (yellows).

Algorithm 1: Overall Training Procedure

NER+PA+RL

Input: Human Annotated (A) + Distantly Labeled Data (D)

- 1 Pre-train NER w/ Partial-CRF (NER+PA) on A+D
 - 2 Apply RL on D
 - 3 Train NER+PA using A + cleaned D
-

they are protein, drug, gene, disease names. Intuitively, given a sentence of the words $X = \{x_1, x_2, \dots, x_n\}$, NER assigns unique tag for each word like $y = \{y_1, y_2, \dots, y_n\}$ from a predefined set of categories $y_i \in \Phi, |\Phi| = k$. Our baseline model is a BiLSTM-CRF architecture (Lample et al., 2016; Habibi et al., 2017). The first layer takes character embeddings for each word sequence and then merge the output vector with the word embedding vector to feed into a second BiLSTM layer. The CRF layer comes on top of the last layer to model the dependencies across output tags and locates the best tag sequence by maximizing the log-probability in following equation:

$$\log(p(y|X)) = \log \frac{e^{s(X,y)}}{\sum_{y' \in Y} e^{s(X,y')}} \quad (1)$$

where

$$s(X, y) = \sum_{i=1}^n P_{i,y_i} + \sum_{i=1}^n T_{y_i,y_{i+1}} \quad (2)$$

and \mathbf{P} is a $k \times n$ output tensor of a linear encoder applied to the last BiLSTM layer where $P_{i,j}$ corresponds to the score of the j^{th} tag of the i^{th} word in a sentence. T is a $(k+2) \times (k+2)$ transition tensor which represents transition probability from i^{th} tag to the j^{th} tag. Two additional tags <BOS> and <EOS> are added at the start and end of a sequence, respectively. In order to infer the final

sequence tags the Viterbi algorithm is employed in the CRF model.

3.2 Partial-CRF layer (PA)

As mentioned above, FN instances constitute a common problem in distantly annotated datasets. It is caused by limited coverage of the knowledge base resource, when some of the entity mentions are not found in the resource and followingly labeled as non-entities ('O'). We follow Tsuboi et al. (2008) and treat the result of distant supervision as a partially annotated dataset where non-entity text spans are annotated as any possible tag. Figure 1 illustrates the annotation of distantly supervised examples using the IOBES labeling scheme that we employ.

Let Y_L denote all the possible tag sequences for a distantly supervised sentence X . Then, the conditional probability of the subset Y_L given X is:

$$p(Y_L|X) = \sum_{y \in Y_L} p(y|X). \quad (3)$$

Extending the original equation of the CRF layer (Eq.1) provides the log-probability for the distantly supervised instance:

$$\log(p(Y_L|X)) = \log \frac{\sum_{y' \in Y_L} e^{s(X,y')}}{\sum_{y' \in Y} e^{s(X,y')}}. \quad (4)$$

Using partial annotation, non-entity text spans are annotated as any possible tag. It gives a chance for non-entity text spans to be considered and scored properly in update version of CRF (Partial CRF) and become a part of the most optimal tag sequence.

3.3 Reinforcement Learning for denoising

The RL agent is designed to determine whether the distantly supervised instance is a true positive or not. There are two main components in RL :

Algorithm 2: Reinforcement learning Algorithm to clean FPs in Distantly Labeled Data (D)

Input: Training dataset (A_{train}) + Distantly Labeled Data (D), Pre-train NER+PA on $A_{train} + D$, Validation dataset (A_{val})

- 1 Initialize θ in policy network
- 2 Initialize s^* as all-zero vector with the same dimension of s_j
- 3 **for** epoch $i = 0 \rightarrow N$ **do**
- 4 **for** instance $d_j \in D$ **do**
- 5 Provide s_j using NER+PA model $\tilde{s}_j = \text{concatenation}(s_j, s^*)$
- 6 Randomly sample $a_j \sim \pi(a; \theta, \tilde{s}_j)$; compute $p_j = \pi(a; \theta, \tilde{s}_j)$, save (a_j, p_j)
- 7 **if** $a_j == 0$ **then**
- 8 save \tilde{s}_j into Ψ_i
- 9 Recompute the s^* as an average of $\forall \tilde{s}_j \in \Psi_i$
- 10 $D_i = D - (\forall d_j; j \in \Psi_i)$
- 11 Train NER+PA on $A_{train} + D_i$
- 12 Calculate F_1^i on A_{val} and save F_1^i and Ψ_i
- 13 $r_i = F_1^i - F_1^{i-1}$
- 14 Find Ω_i, Ω_{i-1} (Eq. 6)
- 15 Update Policy network (Eq. 5)
- 16 Update $D = D - (\forall d_j; j \in \Psi_N)$
- 17 Re-train NER+PA on $A + D$

I) environment II) policy based agent. Following Qin et al. (2018), we model the environment as a Markov Decision Process (MDP), where we add information from the previous state to the current state. The policy based agent is formulated based on the Policy Gradient Algorithm (Sutton et al., 1999), where we update the policy model by computing the reward after finishing the selection process for the whole training set. The algorithm 2 presents additional details of the RL strategy in our NER model. The following subsections describe the elements of the RL agent.

State: The RL agent interacts with the environment to decide about instances at the sentence level. A central component of the environment is the current and previous state in the selection process. The state S_i in step i represents the current instances as well as their label sequences. Following Yang et al. (2018) the state vector S_i includes: I) the vector representation of instances before the Partial-CRF layer, where we concatenate the outputs of the first and last nodes in the BiLSTM layer of the base NER model, and II) the label sequence scores calculated by the linear encoder before the Partial-CRF model. (i.e., $P_{i,j}$ in Eq. 2). If a word is annotated with a certain label, the score will be the corresponding value of the label, otherwise, the score will be the mean of all possible labels of the

word in the linear encoder. These two vectors are concatenated to represent the current state. To satisfy the MDP, the average vector of the removed instances in the earlier step $i - 1$ is concatenated to the current state and represents the state for the RL agent.

Reward: If the RL agent filters out the FP instances from the noisy dataset, the NER model will achieve improved performance. Accordingly, the RL agent will receive a positive reward, otherwise, the agent will received a negative reward. Following Qin et al. (2018), we model the reward as a change of the NER performance; particularly, we adapt the $F1$ score to calculate the reward as the difference between $F1$ scores of the adjacent epochs (i.e., $r_i = F_1^i - F_1^{i-1}$).

Policy Network: The policy network $\pi(a_j; \theta_i, s_j)$ is a feed forward network with two fully-connected hidden layers. It receives the state vector for each distantly supervised instance and then determines whether the instance is a false positive or not. The π as a classifier with parameter θ decides an action $a_j \in \{1, 0\}$ for each $s_j \in S_j$. The loss function for the policy network is formulated based on the policy gradient method (Sutton et al., 1999) and the

REINFORCE algorithm (Williams, 1992). Since we calculate the reward as a difference between $F1$ scores in two contiguous epochs, the agent will be compensated for a set of actions that has direct impact on the performance of the NER model in the current epoch. In other words, the different parts of the removed instances in each epoch are the reason of the change in $F1$ scores. Accordingly, the policy will update using the following gradient:

$$\begin{aligned} \theta = \theta + \mu [\nabla_{\theta} \sum_{\Omega_i} \log \pi(a|S; \theta) r_i \\ + \nabla_{\theta} \sum_{\Omega_{i-1}} \log \pi(a|S; \theta) (-r_i)] \end{aligned} \quad (5)$$

According to Qin et al. (2018), assuming Ψ_i is removed in epoch i :

$$\begin{aligned} \Omega_i &= \Psi_i - (\Psi_i \cap \Psi_{i-1}) \\ \Omega_{i-1} &= \Psi_{i-1} - (\Psi_i \cap \Psi_{i-1}) \end{aligned} \quad (6)$$

This means that if there is an increase in $F1$ at the current epoch i , we will assign a positive reward to the instances that have been removed in epoch i and not in epoch $i - 1$ and negative reward to the instances that have been removed in epoch $i - 1$ and not in the current epoch.

4 Experiments

We perform experiments on four benchmark datasets to compare our method to similar techniques and investigate the impact of the number of available annotated sentences for our approach.

4.1 Experimental Settings

Datasets: Our approach requires an annotated dataset, a knowledge resource and a corpus of raw text. We rely on the resources used by Shang et al. (2018b) and Yang et al. (2018) for English and Chinese, respectively, as well as their train-test splits. For all datasets, we employ a IOBES labeling scheme. Below we briefly describe the datasets:

- BC5CDR is from BioCreative V Chemical Disease Relation task and contains 12,852 'Disease' and 15,935 'Chemical' entity mentions in 1,500 articles. It is already partitioned into a training, a development and a testing set. The related dictionary comes from the MeSH

database² and the CTD chemical and Disease³ vocabularies and contains 322,882 'Disease' and 'Chemical' entities. As a raw text, we use a corpus consisting of 20,217 sentences that is provided in Shang et al. (2018b) and extracted from PubMed papers.

- LaptopReview containing laptop aspect term is taken from the SemEval 2014 Challenge, Task 4 Subtask 1 (Pontiki et al., 2014). The 3,845 review sentences are annotated with 3,012 'AspectTerm' mentions. We extract 15,000 sentences from the Amazon laptop review dataset⁴ as a raw text. Wang et al. (2011) design this dataset for the aspect-based sentiment analysis. Thanks to Shang et al. (2018b), they provide the dictionary of 13,457 computer terms crawled from a public website⁵.
- EC is a Chinese dataset from the e-commerce domain. We choose this dataset in order to compare our results to the approach by Yang et al. (2018). There are 5 entity types: 'Brand', 'Product', 'Model', 'Material' and 'Specification' on user queries. This corpus contain 1,200 training instances, 400 in development set and 800 in test set. Yang et al. (2018) provide the dictionary of 927 entries and 2,500 sentence as a raw text.
- NEWS is another Chinese dataset in the news domain. It is annotated with PERSON type and provided by Yang et al. (2018). The NEWS dataset contains 3,000 sentences as training, 3,328 as dev data, and 3,186 as testing data. Yang et al. (2018) apply distant supervision to raw data and obtain 3,722 annotated sentences.

Pre-trained Embeddings: We employ pre-trained embeddings as initialization for the embedding layer of the LSTM layers. For the biomedical dataset, we use pre-trained 200-dimensional word vectors trained on PubMed abstracts, all PubMed Central (PMC) articles and English Wikipedia (Pyysalo et al., 2013). Standard pre-trained GloVe 100-dimensional word vectors are employed for the LaptopReview dataset. In

²https://www.nlm.nih.gov/mesh/download_mesh.html

³<http://ctdbase.org/downloads/>

⁴<http://times.cs.uiuc.edu/~wang296/Data/>

⁵<https://www.computerhope.com/jargon.htm>

Model	Data	Pr.	Re.	F1
Liu et al. (2017) *	BC5CDR	88.84	85.16	86.96
Wang et al. (2018)		89.10	88.47	88.78
Beltagy et al. (2019)**		-	-	88.94
NER+PA+RL (This work)		92.05	87.91	89.93
Winner system in Pontiki et al. (2014)	LaptopReview	84.80	66.51	74.55
NER+PA+RL (This work)		81.07	74.01	77.38
Yang et al. (2018)	EC	61.57	61.33	61.45
NER+PA+RL (This work)		61.86	65.36	63.56
Yang et al. (2018)	NEWS	81.63	76.95	79.22
NER+PA+RL (This work)		80.20	79.88	80.04

Table 1: NER models comparison. The results on the Chinese EC and NEWS dataset are without high-quality phrases. *: is the base NER model in our approach and results are reported by Wang et al. (2018). **: is the state-of-the-art result on BC5CDR dataset, where they use Pretrained Contextualized Embeddings for Scientific Text (SciBERT) in Ma and Hovy (2016) for NER.

our experiments on the EC dataset, we use the 100-dimensional Chinese character embeddings provided by Yang et al. (2018) and trained on user-generated text.

Evaluation: We report the performance of the model on the test set as the micro-averaged precision, recall and $F1$ score. A predicted entity is counted as a true positive if both the entity boundary and entity type is the same as the ground-truth (i.e., exact match). To alleviate randomness of the scores, the mean of five different runs are reported.

Model Variants: We use slightly different variants of our model for English and Chinese. For English we follow Liu et al. (2017) in leveraging a language model to extract character-level knowledge. We keep the parameters in the model the same as in the original work. In order to compare to state-of-the-art models, we follow the same approach during training (i.e., by merging the training and development data as a training set in BC5CDR and randomly selecting 20% from the training set as the development set in LaptopReview). For the Chinese EC dataset, we only use character-based LSTM and CRF layers and discard the word-based LSTM and language model. For a fair comparison, the model parameters are set to be the same as in Yang et al. (2018). For RL, the batch size, optimizer and learning rate are equal to the parameters in the related NER model. We use 100 epochs in RL and initialize the average vector of the removed sentences as an all-zero vector.

High-Quality Phrases: Considering all non-entity spans (i.e., 'O' type) as a potential entity provides noise in the Partial-CRF process. To address this issue, we use a set of quality multi-word and single-word phrases, provided by Shang et al. (2018b) and obtained using their AutoPhrase method (Shang et al., 2018a). Note that this resource is available only for the English datasets, therefore, it is not included in the experiments on the Chinese datasets. When using these phrases, we assign all possible tags only for the token spans that are matched with this extended list. In our model, we treat the high-quality phrases as potential entities and we assign all possible entity types in annotation of distantly supervised sentences. For example, in Figure 1, we could only find the word 'leprosy' in this list, therefore, in annotation we assign all possible tags to this token and the other non-entity tokens remain as 'O'.

5 Performance Comparison

The first two rows of Table 1 depicts the comparison of the proposed model to the state-of-the-art NER models on the English datasets. We observe that the NER+PA+RL model achieves higher $F1$ scores on the different datasets compared to the other models. In order to compare to the RL based approach in Yang et al. (2018), we run the model without high-quality phrases on the Chinese EC and NEWS datasets. Our design boosts the reported $F1$ score from 61.45 and 79.22 in the original to 63.56 and 80.04 with our model on EC and NEWS datasets, respectively. The experiments on

Model Variant	Data	Pr.	Re.	F1
NER+PA	BC5CDR	85.82	88.58	87.18
NER+PA \oplus		91.28	87.07	89.13
NER+PA+RL		87.00	89.04	88.01
NER+PA+RL \oplus		92.05	87.91	89.93
NER+PA	Laptop Review	61.00	70.80	65.53
NER+PA \oplus		66.36	66.06	66.21
NER+PA+RL		80.47	73.70	76.94
NER+PA+RL \oplus		81.07	74.01	77.38

Table 2: Result with different setting of the distantly supervised NER model. \oplus indicates that we use the list of high-quality phrases along with the dictionary to annotate raw text.

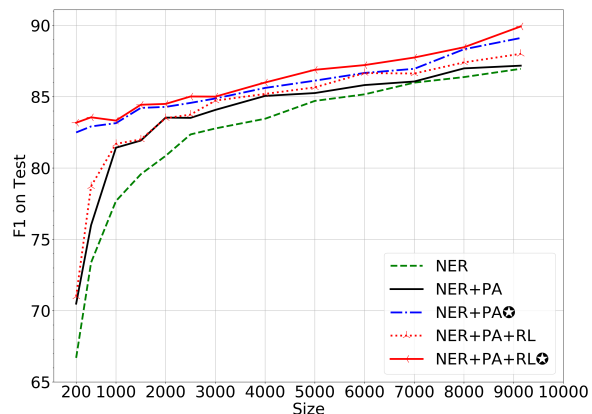


Figure 2: Performance of the different configuration: F1 Score on Test vs, the number of human annotated sentences

the Chinese datasets show that the different design of the RL module leads to improved results.

We further investigate the impact of the different components of the model (Table 2) in the two English datasets via ablation experiments, where we contrast the use of partial annotation (PA) and reinforcement-based denoising RL, with and without the high-quality phrases (\oplus). The experiments confirm the efficiency of the PA and RL modules in resolving FN and FP issues in the distantly labeled dataset. The results also corroborate Shang et al. (2018b) in showing that incorporation of the high-quality phrases always leads to a boost in the precision and subsequently in F1 score.

6 Size Of Gold Dataset

In all the previous experiments, we take advantage of the availability of an annotated dataset. However, one of the challenges in domain specific NER is the availability of a gold supervision data. We here examine the performance of

Method	Data	Pr.	Re.	F1
Dictionary Match	BC5CDR	93.93	58.35	71.98
Fries et al. (2017)		84.98	83.49	84.23
Shang et al. (2018b)		88.96	81.00	84.80
NER+PA+RL \oplus		88.73	77.51	82.74
Dictionary Match	Laptop Review	90.68	44.65	59.84
Giannakopoulos et al. (2017)		74.51	31.41	44.37
Shang et al. (2018b)		72.27	59.79	65.44
NER+PA+RL \oplus		68.63	56.88	62.21

Table 3: Unsupervised NER Performance Comparison. The proposed method is trained only on distantly labeled data.

the proposed model on the BC5CDR corpus by selecting increasing amounts of annotated instances from the gold dataset. As shown in Figure 2, the proposed method achieves a performance of 83.18 only with 2% of the annotated dataset. Whereas the base NER model, requires almost 45% of the ground truth sentences to reach the same performance. This indicates that with a small set of human annotated data, our model can deliver relatively good performance.

We also carry out experiments on the BC5CDR and LaptopReview test sets, where our model is trained exclusively on distantly annotated data. We report the outcome together with the scores of the other state-of-the-art unsupervised methods in Table 3, where we also compare to simple dictionary matching. It is clear that the model of Shang et al. (2018b) (AutoNER) is still the best performing NER method on BC5CDR and LaptopReview datasets in an unsupervised setup. However, as is clear from Figures 3-a and 3-c in Shang et al. (2018b)), if there is at least some manually labeled data available, our method makes better use of the gold supervision compared to the AutoNER system in the similar training scenario. It is also worth noting that the approach proposed by Fries et al. (2017) utilizes extra human effort to design regular expressions and requires specialized hand-tuning.

7 Conclusion and Future work

This work presents an approach to alleviate the problems of auto-generated data in NER. The performance-driven, policy-based reinforcement learning module removes the sentences with FPs, whereas the adapted Partial-CRF layer deals with FNs. We examine the impact of each component in ablation experiments. Combining these in a su-

pervised setting leads to state-of-the-art results on three benchmark datasets from different domains and different languages.

Future work will extend the study to improve the performance of the model in unsupervised fashion and extend our study to additional domains and languages.

References

- Alan Akbik, Duncan Blythe, and Roland Vollgraf. 2018. [Contextual string embeddings for sequence labeling](#). In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1638–1649, Santa Fe, New Mexico, USA. Association for Computational Linguistics.
- Isabelle Augenstein, Diana Maynard, and Fabio Ciravegna. 2014. Relation extraction from the web using distant supervision. In *EKAW*, volume 8876 of *Lecture Notes in Computer Science*, pages 26–41. Springer.
- Iz Beltagy, Arman Cohan, and Kyle Lo. 2019. [Scibert: Pretrained contextualized embeddings for scientific text](#). *CoRR*, abs/1903.10676.
- Jun Feng, Minlie Huang, Li Zhao, Yang Yang, and Xiaoyan Zhu. 2018. Reinforcement learning for relation classification from noisy data. *CoRR*, abs/1808.08013.
- Jason A. Fries, Sen Wu, Alexander Ratner, and Christopher Ré. 2017. [Swelshark: A generative model for biomedical named entity recognition without labeled data](#). *CoRR*, abs/1704.06360.
- Athanasios Giannakopoulos, Claudiu Musat, Andreea Hossmann, and Michael Baeriswyl. 2017. [Unsupervised aspect term extraction with b-LSTM & CRF using automatically labelled datasets](#). In *Proceedings of the 8th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*, pages 180–188, Copenhagen, Denmark. Association for Computational Linguistics.
- Maryam Habibi, Leon Weber, Mariana L. Neves, David Luis Wiegandt, and Ulf Leser. 2017. Deep learning with word embeddings improves biomedical named entity recognition. In *Bioinformatics*.
- Guillaume Lample, Miguel Ballesteros, Sandeep Subramanian, Kazuya Kawakami, and Chris Dyer. 2016. [Neural architectures for named entity recognition](#). *CoRR*, abs/1603.01360.
- Liyuan Liu, Jingbo Shang, Frank F. Xu, Xiang Ren, Huan Gui, Jian Peng, and Jiawei Han. 2017. [Empower sequence labeling with task-aware neural language model](#). *CoRR*, abs/1709.04109.
- Xuezhe Ma and Eduard Hovy. 2016. [End-to-end sequence labeling via bi-directional LSTM-CNNs-CRF](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1064–1074, Berlin, Germany. Association for Computational Linguistics.
- Mike Mintz, Steven Bills, Rion Snow, and Dan Jurafsky. 2009. [Distant supervision for relation extraction without labeled data](#). In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2 - Volume 2*, ACL '09, pages 1003–1011, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Alexandre Passos, Vineet Kumar, and Andrew McCallum. 2014. [Lexicon infused phrase embeddings for named entity resolution](#). *CoRR*, abs/1404.5367.
- Matthew E. Peters, Mark Neumann, Mohit Iyyer, Matt Gardner, Christopher Clark, Kenton Lee, and Luke Zettlemoyer. 2018. [Deep contextualized word representations](#). *CoRR*, abs/1802.05365.
- Maria Pontiki, Dimitris Galanis, John Pavlopoulos, Harris Papageorgiou, Ion Androutsopoulos, and Suresh Manandhar. 2014. [SemEval-2014 task 4: Aspect based sentiment analysis](#). In *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, pages 27–35, Dublin, Ireland. Association for Computational Linguistics.
- S. Pyysalo, F. Ginter, H. Moen, T. Salakoski, and S. Ananiadou. 2013. [Distributional semantics resources for biomedical text processing](#). In *Proceedings of LBM 2013*, pages 39–44.
- Pengda Qin, Weiran Xu, and William Yang Wang. 2018. [Robust distant supervision relation extraction via deep reinforcement learning](#). *CoRR*, abs/1805.09927.
- Lev Ratinov and Dan Roth. 2009. [Design challenges and misconceptions in named entity recognition](#). In *Proceedings of the Thirteenth Conference on Computational Natural Language Learning*, CoNLL '09, pages 147–155, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Xiang Ren, Ahmed El-Kishky, Chi Wang, Fangbo Tao, Clare R. Voss, and Jiawei Han. 2015. [Clustype: Effective entity recognition and typing by relation phrase-based clustering](#). In *KDD*, pages 995–1004. ACM.
- Sebastian Riedel, Limin Yao, and Andrew McCallum. 2010. [Modeling relations and their mentions without labeled text](#). In *Proceedings of the 2010 European Conference on Machine Learning and Knowledge Discovery in Databases: Part III*, ECML PKDD'10, pages 148–163, Berlin, Heidelberg. Springer-Verlag.
- Jingbo Shang, Jialu Liu, Meng Jiang, Xiang Ren, Clare R Voss, and Jiawei Han. 2018a. Automated phrase mining from massive text corpora. *IEEE Transactions on Knowledge and Data Engineering*.

- Jingbo Shang, Liyuan Liu, Xiang Ren, Xiaotao Gu, Teng Ren, and Jiawei Han. 2018b. Learning named entity tagger using domain-specific dictionary. In *EMNLP*.
- Richard S. Sutton, David McAllester, Satinder Singh, and Yishay Mansour. 1999. Policy gradient methods for reinforcement learning with function approximation. In *Proceedings of the 12th International Conference on Neural Information Processing Systems, NIPS'99*, pages 1057–1063, Cambridge, MA, USA. MIT Press.
- Yuta Tsuboi, Hisashi Kashima, Shinsuke Mori, Hiroki Oda, and Yuji Matsumoto. 2008. Training conditional random fields using incomplete annotations. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008)*, pages 897–904, Manchester, UK. Coling 2008 Organizing Committee.
- Hongning Wang, Yue Lu, and ChengXiang Zhai. 2011. Latent aspect rating analysis without aspect keyword supervision. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '11*, pages 618–626, New York, NY, USA. ACM.
- Xuan Wang, Yu Zhang, Xiang Ren, Yuhao Zhang, Marinka Zitnik, Jingbo Shang, Curtis Langlotz, and Jiawei Han. 2018. Cross-type biomedical named entity recognition with deep multi-task learning. *CoRR*, abs/1801.09851.
- Ronald J. Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, 8(3):229–256.
- Yaosheng Yang, Wenliang Chen, Zhenghua Li, Zhengqiu He, and Min Zhang. 2018. Distantly supervised NER with partial annotation learning and reinforcement learning. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 2159–2169, Santa Fe, New Mexico, USA. Association for Computational Linguistics.