

Exploiting Noisy Data in Distant Supervision Relation Classification

Kaijia Yang, Liang He, Xin-yu Dai, Shujian Huang, Jiajun Chen

National Key Laboratory for Novel Software Technology,

Nanjing University, Nanjing, 210023, China

{yangkj, heliang}@nlp.nju.edu.cn

{daixinyu, huangsj, chenjj}@nju.edu.cn

Abstract

Distant supervision has obtained great progress on relation classification task. However, it still suffers from noisy labeling problem. Different from previous works that underutilize noisy data which inherently characterize the property of classification, in this paper, we propose RCEND, a novel framework to enhance Relation Classification by Exploiting Noisy Data. First, an instance discriminator with reinforcement learning is designed to split the noisy data into correctly labeled data and incorrectly labeled data. Second, we learn a robust relation classifier in semi-supervised learning way, whereby the correctly and incorrectly labeled data are treated as labeled and unlabeled data respectively. The experimental results show that our method outperforms the state-of-the-art models.

1 Introduction

Relation classification plays a crucial role in natural language processing (NLP) tasks, such as question answering and knowledge base completion (Xu et al., 2016; Han et al., 2018a). The goal of relation classification is to predict relations of the target entity pair given a plain text. Traditional supervised learning methods (Zelenko et al., 2002; Bunescu and Mooney, 2005; Zhou et al., 2005) heavily rely on large scale annotated data which is time and labor consuming. Mintz et al. (2009) proposed distant supervision (DS) to automatically generate training data for relation classification based on the assumption that if two target entities have a relation in knowledge base (KB), sentences containing this entity pair might express the relation. For example, if a relational fact $\langle \text{Apple}, \text{founder}, \text{Steve Jobs} \rangle$ exists in KB, distant supervision will assign *founder* as the label of all sentences that contain “Apple” and “Steve Jobs” together.

Sentence	DS	Gold
S1: Al Gore was waiting to board a commercial flight from Nashville to Miami...	LivedIn	NA
S2: There were also performers who were born in Louisiana , including Lucinda Williams ...	LivedIn	BornIn
S3: Boggs was married, had three young children and lived in Brewster	NA	LivedIn

Table 1: Examples of noisy labeling problem in distant supervision relation classification. S1 and S2 are heuristically labeled as *LivedIn* by DS, but neither of them mention the relation while S2 mentions the *BornIn* relation. S3 expresses the *LivedIn* relation but it is mislabeled as *NA* since no relation of the entity pair exist in KB.

However, it suffers from noisy labeling problem due to the irrelevance of aligned text and incompleteness of KB, which consists of **false positives** and **false negatives**. The false positives means that not all sentences containing two entities mention the relation in KB, such as S1 and S2 in Table 1. And the false negatives are sentences are mislabeled as no relation (*NA*) due to the absence of relational fact in KB even though they express the target relation, such as S3 in Table 1.

In order to reduce the impact of noisy data, previous works (Riedel et al., 2010; Hoffmann et al., 2011; Surdeanu et al., 2012; Zeng et al., 2015; Lin et al., 2016; Han et al., 2018b) adopt Multi-Instance Learning (MIL) for relation classification. Recent studies (Feng et al., 2018; Qin et al., 2018b,a) introduce reinforcement learning (RL) and adversarial learning to filter out incorrectly labeled sentences and achieve significant improvements. However, there are two remaining challenges of noisy labeling problem.

- Most of these approaches focus on solving the false positives but overlook false nega-

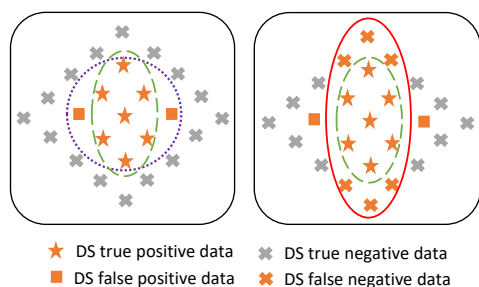


Figure 1: Illustration of false positive and false negative cases

tives. As illustrated in Figure 1, they concentrate on discovering the false positive instances¹ which are suppressed or removed at last and obtain a better decision boundary (green dashed line) than without consideration of false positive instances. Nevertheless, there are still a lot of false negative instances expressing similar semantic information with positive data. These instances also provide evidence for the target relation. The incorrect labels will weaken the discriminative capability of available features and confuse the model if they stay the same. However, when we remedy the label correctly, we indeed possess the optimal decision boundary (red solid line).

- There lacks an effective method to fully utilize noisy data of distant supervision. (Xu et al., 2013; Liu et al., 2017) apply methods such as pseudo-labels to directly correct the label of noisy data and Luo et al. (2017) design a dynamic transition matrix to model noise patterns. They still suffer from the drawback of error propagation during training.

To tackle the above challenges, we propose a novel framework exploiting noisy data to enhance distant supervision relation classification. We design an instance discriminator with reinforcement learning to recognize both false positive and false negative instances simultaneously, and further split the noisy dataset into two sets, representing correctly labeled and incorrectly labeled data respectively. Additionally, we learn a robust relation classifier applying a semi-supervised learning method, whereby the correctly and incorrectly labeled data are regarded as labeled and unlabeled data. On the one hand, we mitigate the

¹In this paper, instance is the same as sentence

side effect of incorrectly labeled data by recognizing them and treating them as unlabeled data. On the other hand, taking full advantage of the incorrectly labeled data in semi-supervised learning way facilitates robust property of model and improves generalization performance. Our contributions in this work are three-fold:

- We propose a deep reinforcement learning framework to discriminate both false-positive and false-negative instances simultaneously.
- We introduce a semi-supervised learning method to fully exploit the noisy data in distant supervision relation classification.
- We conduct experiments on a widely used benchmark dataset and the results show that our method achieves significant improvements as compared with strong baselines.

2 Related work

Many efforts based on supervised learning (Zelenko et al., 2002; Bunescu and Mooney, 2005; Zhou et al., 2005) have been devoted to relation classification. As is well-known, achieving a good performance while applying supervised learning paradigm requires a large amount of high-quality annotated data. To address the issue of data sparsity, Mintz et al. (2009) propose distant supervision to automatically annotate large scale training data, which inevitably results in noisy labeling problem.

To tolerate noisy instances in positive examples, most early approaches employ multi-instance learning framework, including multi-instance single-label learning (Riedel et al., 2010) and multi-instance multi-label learning (Hoffmann et al., 2011; Surdeanu et al., 2012). Recently, deep learning has also been introduced to propose an end-to-end convolutional neural network for relation classification (Zeng et al., 2014). In the sentences bag of one entity pair, Zeng et al. (2015) select the most reliable sentence, and Lin et al. (2016) propose attention schemes to de-emphasize unreliable sentences. Han et al. (2018b) incorporate hierarchical information of relations to enhance the attention scheme. But they fail to handle the issue where all sentences in one bag are mislabeled.

Feng et al. (2018); Qin et al. (2018b,a) further achieve improvement by using reinforcement

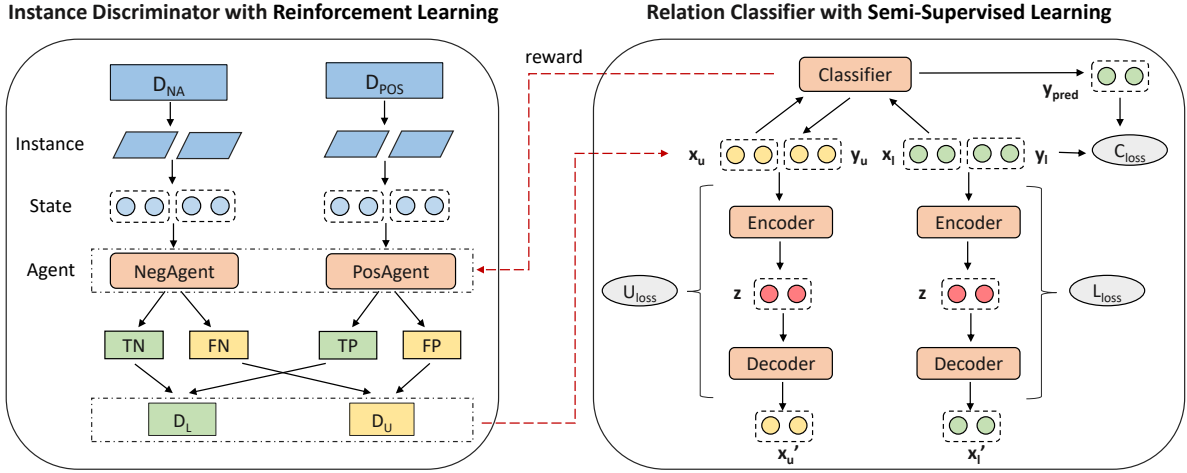


Figure 2: The framework of train process. Instance discriminator, consisting of PosAgent and NegAgent, aims to recognize false-positive (FP) and false-negative (FN) instances from positive dataset (\mathcal{D}_{POS}) and NA dataset (\mathcal{D}_{NA}) respectively. Afterward, true-positive (TP) and true-negative (TN) instances are split into labeled data (\mathcal{D}_l) while FP and FN instances are split into unlabeled data (\mathcal{D}_u). We adopt SemiVAE, which consists of an encoder, a decoder and a classifier, to train a robust relation classifier with semi-supervised learning utilizing the above data $\mathcal{D}_l, \mathcal{D}_u$. More details are introduced in Section 3.2 and 3.3.

learning and adversarial learning to explicitly remove incorrectly labeled sentences. However, they neglect the useful inherent information of those sentences which should be replaced correctly. In other words, they remove the noise rather than utilize it in the right way.

Furthermore, Xu et al. (2013) correct false negative instances by using pseudo-relevance feedback to expand the origin knowledge base. Liu et al. (2017) apply a dynamic soft-label instead of the immutable hard label produced by DS during the training process. Luo et al. (2017) design a transition matrix which characterizes the underlying noise pattern to correct noisy labels. They utilize the noisy data and address the false negative problem to some extent, but they still suffer from the drawback that errors may be propagated because the model is unable to correct its own mistakes.

In this work, we propose a unified framework for learning a discriminator to recognize both false-positive and false-negative instances with reinforcement learning, and utilizing the incorrectly labeled data as unlabeled data in semi-supervised learning way.

3 Methodology

In this section, we introduce our framework and the details of instance discriminator and relation classifier as follows.

3.1 Framework

In MIL paradigm, the entire instances are split into multiple entity-pair bags $\{B_{h_i, t_i}\}_{i=1}^k$. The sentences in $B_{h, t}$ mention both head entity h and tail entity t . Here we denote dataset as $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^n$, where x_i is a sentence associated with the corresponding entity pair, y_i is a noisy relation label produced by distant supervision and n is the total number of sentences contained in each bag. As mentioned above, NA is a special relation which indicates the sentence does not express any relations in the KB. We define other relations in the KB as positive relations. Accordingly, we split the dataset into \mathcal{D}_{POS} and \mathcal{D}_{NA} .

In the scenario of distant supervision, an ideal model is not only capable of capturing valid supervision information about correctly labeled data with less noise, but also leveraging information contained in incorrectly labeled data by correcting the noisy label implicitly.

As a result, we solve the task of distant supervision relation classification in two steps. As depicted in Figure 2, we design an instance discriminator to heuristically recognize false positive and false negative instances from the noisy distantly-supervised dataset with reinforcement learning. The correctly labeled instances discovered by the discriminator are split into labeled data while the incorrectly labeled ones are split into unlabeled data. The details of the instance discriminator

will be introduced in Section 3.2. After scanning the entire noisy dataset, we train a robust classifier with semi-supervised learning utilizing above labeled data and unlabeled data. The details of the relation classifier will be introduced in Section 3.3. Meanwhile, the relation classifier provides rewards to the instance discriminator for updating parameters of its policy function.

3.2 Instance discriminator

We regard recognizing incorrectly labeled instances as a reinforcement learning problem. The instance discriminator acts as an agent interacting with the environment that consists of a noisy dataset and a relation classifier. The agent is parameterized with a policy network $\pi(a|s; \theta)$ which gives the probability distribution of action a at each state s and receives reward r from the relation classifier to update parameters θ . Note that *NA* indicates that there is no relation between two entities or that the relation is of no interest. The relation *NA* is very ambiguous since instances have no unified pattern. Thus we cannot decide whether a sentence belongs to *NA* only by the fact that it does not express any other positive relation. Under this consideration, we adopt two agents, PosAgent and NegAgent, to recognize false positive and false negative instances respectively. The definitions of the components in RL are introduced as follows.

State The state includes the semantic and syntactic information of current sentence and the relation label given by DS. We use a piecewise convolutional neural network (PCNN) (Zeng et al., 2015) to convert each sentence into real-valued vector x and build a class representation matrix M to represent each relation type. As we decide whether the current sentence is correctly labeled according to the similarity between the semantic meanings of sentence and relation, we only take the current sentence into consideration without sentences in early states. For PosAgent, state s_p is the concatenation of the current sentence vector x and corresponding relation embedding. For NegAgent, we represent state s_n by the vector of relational scores based on the representation of the current sentence x .

$$\begin{aligned} s_p &= [x; M_{[y]}] \\ s_n &= Mx + b \end{aligned} \quad (1)$$

where y is relation label of the current sentence. $b \in \mathbb{R}^{n_r}$ is a bias vector and n_r is the number of class.

Action We desire the agent to distinguish whether the current sentence is mislabeled or not. Therefore, the action of our agent is defined as $a_i \in \{0, 1\}$, where 0 indicates the sentence is incorrectly labeled and 1 indicates the sentence is correctly labeled.

Reward The reward function can reflect the advantage of redistributing the noisy data. As previously mentioned, the actions of our agent redistribute noisy data into labeled data and unlabeled data, corresponding to correctly labeled and incorrectly labeled instances. Therefore, the average likelihood of labeled data will be larger than unlabeled data when the agent makes correct actions. We define the difference of likelihood between them as the reward to evaluate the performance of our policy. Consequently, the reward is defined as follows:

$$r = \lambda \left(\frac{1}{|L|} \sum_{x \in L} p_\phi(y|x) - \frac{1}{|U|} \sum_{x \in U} p_\phi(y|x) \right) \quad (2)$$

where L and U is the subset of labeled data and unlabeled data respectively, and y is the relation label given by DS. $p_\phi(y|x)$ is calculated by the relation classifier from the semi-supervised learning framework. λ is used to scale the difference to a rational numeric range.

Training the Policy-based Agent

The objective of the agent is to maximize the expected reward of the actions sampled from the probability distribution. Given a mini-batch data B , following the policy, our agent produces a set of probability distributions of actions $\pi(a_i|s_i; \theta)$. Based on the actions, the agent achieves a performance-driven reward r . We use a policy gradient strategy to compute the gradient and update our agent referring to the policy gradient theorem (Sutton et al., 1999) and the REINFORCE algorithm (Williams, 1992). The parameters of the policy network are updated according to the following gradient:

$$\theta \leftarrow \theta + \alpha \sum_{i=1}^{|B|} \nabla_{\theta} r \log \pi(a_i|s_i; \theta) \quad (3)$$

As the goal of our agent is to determine whether an annotated sentence expresses target relation

with weak supervision, we need a relation classifier to compute the reward for updating the policy network. We first pre-train our classifier on the entire dataset with supervised learning until rough convergence. Then we pre-train the policy network by receiving rewards from the pre-trained classifier with the parameters frozen. The pre-training strategy is necessary as it saves time that would otherwise be spent training the model by trial and error. It is also widely used by other related works (Silver et al., 2016; Bahdanau et al., 2016). The training procedure for instance discriminator is summarized in Algorithm 1.

3.3 Relation Classifier

In order to reach the maximum utilization of noisy data, we train our relation classifier with semi-supervised learning. Below, we introduce PCNN and SemiVAE, the method we adopt for semi-supervised learning.

PCNN

We take the widely used CNN architecture (PCNN) (Zeng et al., 2015; Lin et al., 2016) to encode input sentences into low-dimensional vectors and predict their corresponding relation labels.

Given a sentence containing an entity pair, we represent the i -th word as v_i by concatenating its word embedding w_i and position embedding p_i which encodes the relative distances from it to two entities ($v_i \in \mathbb{R}^d, w_i \in \mathbb{R}^{d_w}, p_i \in \mathbb{R}^{d_p}, d = d_w + d_p$).

Afterward, the convolution layer applies a kernel of the window size l to slide over the input sequence $\{v_1, v_2, \dots, v_m\}$ and output the d_h -dimensional hidden embeddings \mathbf{h} , where $\mathbf{h} \in \mathbb{R}^{m \times d_c}$ and d_c is the number of feature maps.

Then, piecewise max-pooling is used to divide the hidden embeddings into three parts $\{\mathbf{h}^1, \mathbf{h}^2, \mathbf{h}^3\}$ by the position of head and tail entities. We perform max-pooling on each part respectively and get final embedding \mathbf{x} by concatenating the pooling results, where $\mathbf{x} \in \mathbb{R}^{d_s}$ ($d_s = d_c \times 3$).

$$\mathbf{x} = [\max(\mathbf{h}^1); \max(\mathbf{h}^2); \max(\mathbf{h}^3)] \quad (4)$$

Finally, we formalize the probability of predicting y given sentence \mathbf{x} as follows:

$$\begin{aligned} \mathbf{o} &= \mathbf{M}\mathbf{x} + \mathbf{b}, \\ p_\phi(y|\mathbf{x}) &= \frac{\exp(o_y)}{\sum_{k=1}^{n_r} \exp(o_k)} \end{aligned} \quad (5)$$

where $\mathbf{M} \in \mathbb{R}^{n_r \times d_s}$ is the class embeddings of each relation and $\mathbf{b} \in \mathbb{R}^{n_r}$ is a bias vector.

Semi-supervised VAE

SemiVAE, a semi-supervised learning method based on variational inference, is introduced and developed by (Kingma et al., 2014; Xu et al., 2017). The inference model consists of three components as follows. An encoder network $p_\varphi(z|x, y)$ encodes data x and label y into a latent variable z . The decoder network $p_\psi(x|z, y)$ is used to estimate the probability of generating x given z and categorical label y . Finally, classifier $p_\phi(y|x)$ predict the corresponding label y of x . We model both encoder and decoder by multilayer perceptron (MLP) and employ the PCNN model as the classifier in SemiVAE.

For the case of labeled data (x_l, y_l) , the evidence lower bound is:

$$\begin{aligned} \log p_\psi(x_l, y_l) &\geq \mathbb{E}_{p_\varphi(z|x_l, y_l)} [\log p_\psi(x_l|y_l, z)] \\ &\quad + \log p_\psi(y_l) - D_{KL}(p_\varphi(z|x_l, y_l)||p(z)) \quad (6) \\ &= -\mathcal{L}(x_l, y_l) \end{aligned}$$

where first term represent the expectation of the conditional log-likelihood on latent variable z and the last term is Kullback-Leibler divergence between the prior distribution $p(z)$ and the latent posterior $p_\phi(z|x_l, y_l)$.

For the case of unlabeled data x_u , the unobserved label y_u is obtained from the classifier in the inference model. The variational lower bound is:

$$\begin{aligned} \log p_\psi(x_u) &\geq \sum_y p_\phi(y_u|x_u) (-\mathcal{L}(x_u, y_u) + \mathcal{H}(p_\phi(y_u|x_u))) \\ &= -\mathcal{U}(x_u) \end{aligned} \quad (7)$$

where \mathcal{H} denotes the entropy of $p_\phi(y_u|x_u)$.

Since the classifier $p_\phi(y|x)$ is unable to learn directly from labeled data, a classification loss is introduced as:

$$\mathcal{C} = \mathbb{E}_{(x, y) \in \mathcal{D}_l} [-\log p_\phi(y|x)] \quad (8)$$

To maximize the evidence lower bound of both labeled data and unlabeled data and minimize the classification loss, the objective is defined as:

$$\mathcal{J} = \sum_{(x, y) \in \mathcal{D}_l} \mathcal{L}(x, y) + \sum_{x \in \mathcal{D}_u} \mathcal{U}(x) + \beta \mathcal{C} \quad (9)$$

where \mathcal{D}_l and \mathcal{D}_u are labeled and unlabeled data respectively, β is a factor used to scale the classification loss of labeled data.

Algorithm 1 Reinforcement Learning Algorithm for Instance Discriminator.

Input: Origin dataset $\mathcal{D}_{\text{POS}} = \{(x_i, y_i)\}_{i=1}^n$. (To be clear, we demonstrate the training procedure of PosAgent. NegAgent is trained in the same way.)

Output: Labeled data \mathcal{D}_l , unlabeled data \mathcal{D}_u .

- 1: Initialize parameters of policy network as θ .
 - 2: **for** step $t = 1 \rightarrow T$ **do**
 - 3: Partition \mathcal{D}_{POS} into minibatches of size b_s
 - 4: **for** minibatch $B \subset \mathcal{D}_{\text{POS}}$ **do**
 - 5: $B = \{(x_j, y_j)\}_{j=1}^{b_s}$
 - 6: Sample actions for each sentence in B :
 $a_j \sim \pi(a_j | s_j; \theta)$
 - 7: **if** $a_j == 0$ **then**
 - 8: Add x_j to \mathcal{D}_u
 - 9: **else**
 - 10: Add (x_j, y_j) to \mathcal{D}_l
 - 11: **end if**
 - 12: Calculate reward r by Eq.(2)
 - 13: Update θ by Eq.(3)
 - 14: **end for**
 - 15: **end for**
-

Algorithm 2 Semi-supervised Learning Algorithm for Relation Classifier.

Input: Labeled data \mathcal{D}_l , unlabeled data \mathcal{D}_u .

- 1: Initialize parameters of relation classifier as ϕ .
 - 2: **for** epoch $i = 1 \rightarrow N$ **do**
 - 3: Sample m data pair (x_l, y_l) from \mathcal{D}_l
 - 4: Sample m data x_u from \mathcal{D}_u and predict their corresponding unobserved label y_u via $p_\phi(y|x)$
 - 5: Update ϕ by Eq.(9)
 - 6: **end for**
-

After our reinforcement learning process, we obtain an instance discriminator which possesses the capability of recognizing incorrectly labeled instances from the noisy dataset. Additionally, the entire DS dataset \mathcal{D} is split into labeled data \mathcal{D}_l and unlabeled data \mathcal{D}_u . Therefore, we utilize the above data to train SemiVAE model and obtain a robust relation classifier which explicitly learns from correctly labeled data and correct incorrectly labeled data implicitly. The training procedure for relation classifier is summarized in Algorithm 2.

4 Experiment

4.1 Datasets and Evaluation

We evaluate our model on a widely used dataset that is generated by aligning entity pairs from

Batch size b_s	160
Word Dimension d_w	50
Position Dimension d_p	5×2
Convolution Filter Dimension d_c	230
Convolution Window Size l	3
Latent Variable Dimension d_z	100
Dropout p	0.5
Regulator λ, β	100, 2

Table 2: Hyperparameter settings

Freebase with New York Times corpus(NYT)² and developed by (Riedel et al., 2010). Entity mentions are recognized by the Stanford named entity recognizer (Finkel et al., 2005). The relation facts in Freebase are divided into two parts for training and testing respectively. The sentences from the corpus of the years 2005-2006 are used as the training instances, and sentences from 2007 are used as the testing instances. There are 52 positive relations and a special relation *NA*.

Following previous works, we evaluate our model on the held-out evaluation, which compares relation facts extracted from the test corpus with those in Freebase. We adopt aggregated precision/recall curves and precision@N (P@N) to illustrate the performance of our model.

4.2 Parameter Settings

We adopt the Adam (Kingma and Ba, 2014) optimizer to optimize our instance discriminator and relation classifier with learning rate 0.0001 and 0.001 respectively. We also apply dropout to prevent overfitting. More detailed hyperparameter settings are presented in Table 2.

4.3 Overall Evaluations Results

We adopt the following baselines with which we compare our model:

- **Mintz** (Mintz et al., 2009) is the original distantly supervised model. **MultiR** (Hoffmann et al., 2011) and **MIML**(Surdeanu et al., 2012) handle overlapping relation problem with graphical model in multi-instance and multi-instance multi-label framework. All above models are based on handcrafted feature.
- **PCNN+ONE** (Zeng et al., 2015) and **PCNN+ATT** (Lin et al., 2016) are both robust models to solve noisy labeling problem

²<http://iesl.cs.umass.edu/riedel/ecml/>

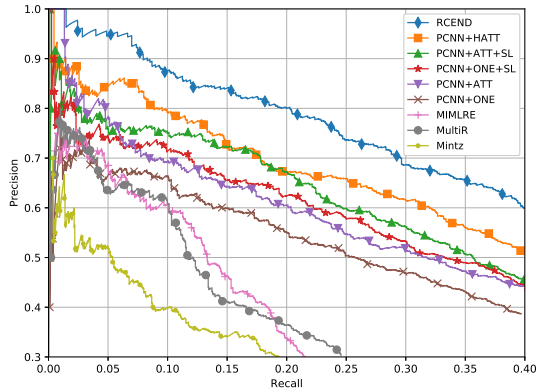


Figure 3: Precision-recall curves of our model and baselines.

P@N	100	200	300	Mean
PCNN+ONE	72.3	69.7	64.1	68.7
PCNN+ATT	76.2	73.1	67.4	72.2
PCNN+ONE+SL	84.0	81.0	74.0	79.7
PCNN+ATT+SL	87.0	84.5	77.0	82.8
PCNN+HATT	88.0	79.5	75.3	80.9
RCEND	95.0	87.5	84.4	88.9

Table 3: Top-N precision (P@N) of our model and baselines

based on the at-least-one assumption and selective attention. **PCNN+HATT** (Han et al., 2018b) is a attention-based method which employs hierarchical attention to exploit correlations among relations.

- **PCNN+ONE+SL** and **PCNN+ATT+SL** (Liu et al., 2017) use a soft-label method to alleviate the negative impact of the noisy labeling problem.

We compare our model with aforementioned baselines and the results are shown in Figure 3. From the overall result we can see that: (1) All feature-based models perform poorly as their features are derived from NLP tools, which will generate errors that propagate through in model. (2) PCNN+ONE and PCNN+ATT boost the performance because they reduce noise in the bag of entity pair by selecting the most confident sentence or de-emphasize the incorrectly labeled sentences with an attention mechanism. (3) When PCNN+ONE and PCNN+ATT use soft labels, they achieve an improvement. This indicates correcting the noisy label is helpful to relation classification in MIL scheme. (4) PCNN+HATT further enhances the performance as it incorporates hierarchical information of relations to improve the attention mechanism. (5) Our method RCEND achieves the best precision over the entire recall

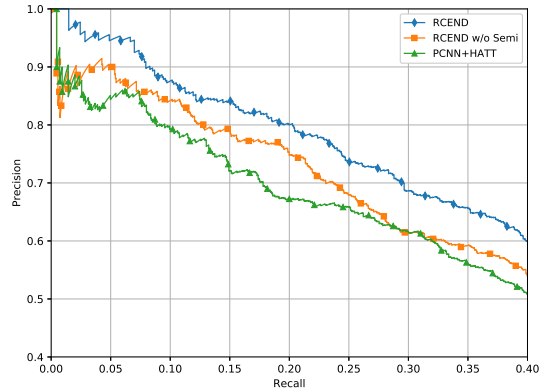


Figure 4: Precision-recall curves of our model with different settings.

P@N	100	200	300	Mean
RCEND	95.0	87.5	84.4	88.9
RCEND w/o Semi	90.0	84.6	79.1	84.6
RCEND(P)	87.1	82.1	80.1	83.3
RCEND(N)	89.1	85.1	81.1	85.1

Table 4: Top-N precision (P@N) of our model with different settings.

range compared with all baselines. The performance achieves further improvement when we regard the incorrectly labeled sentences as unlabeled data and adopt a semi-supervised learning method to train our model. It shows that exploiting noisy data with our method is beneficial to promote distant supervision relation classification.

We also report the result of Precisions@N (100, 200, 300) in Table 3. We can see that our method outperforms the baselines on the precision values of top N triples extracted.

4.4 Impact of Unlabeled Data

To further verify the impact of the unlabeled data, we conduct experiments with both utilization and non-utilization of unlabeled data. The results are presented in Figure 4. Note that, the method RCEND w/o Semi is similar to the method proposed by (Feng et al., 2018), which only removes the incorrectly labeled sentences but does not fully utilize them. We can see that it achieves higher precision over the entire level of recall compared to PCNN+HATT, the best noise-tolerate method in MIL scheme, which shows that removing noise is better than dealing with them with soft attention weights. However, it is still unable to surpass our method. In Table 4, our method also shows notable improvement over RCEND w/o Semi. This demonstrates that fully utilizing noisy data is more advantageous than reducing them.

Type	Sentence	Predict	DS
FN	C1: [Oliver O’Grady] is now a silver-haired , twinkly-eyed resident of [Ireland] , where Ms. Berg often films him in parks ...	Nationality	NA
	C2: ... said [John Allison] , editor of [Opera] magazine , based in London.	EmployedBy	NA
	C3: [Jean-Pierre Bacri] is a famous writer , who is too self-centered to care about his lonely , overweight , 20-year-old daughter , [Lolita Marilou Berry]...	ChildrenOf	NA
FP	C4: they wanted to interview [Bill Cosby] after they met with a former Temple University employee who has accused him of groping her in his home in suburban [Philadelphia]	LivedIn	BornIn
	C5: “Without the fog , [London] wouldn’t be a beautiful city.”the French painter Claude Monet wrote to his wife , Alice , during one of his long visits to [England] from France.	NA	Capital
	C6: MTV Goes to Africa MTV opened its first local music channel in Africa this week , a step touted by the singer [Lebo Mathosa] , above , at an MTV event in [Johannesburg].	NA	DieIn

Table 5: Examples for case study. The first three sentences are examples of false negative case and the final three are examples of false positive case.

This can be partially explained due to the label rectification of the incorrectly labeled data during semi-supervised learning with correctly labeled data which improves the generalization performance.

4.5 Impact of False Positives and False Negatives

The goal of this experiment is to inspect whether the relation classifier is enhanced more through the utilization of false negatives or through the utilization of false positives. As depicted in Figure 5, RCEND(P) only recognizes the false positive sentences in \mathcal{D}_{POS} by PosAgent and regards them as unlabeled data. Likewise, RCEND(N) only discovers and utilizes false negative sentences. RCEND(P) and RCEND(N) behave similarly and achieve further improvement when utilizing both false-positive and false-negative sentences, which implies that both of them are important and promote the ability of our relation classifier. And the results in Table 4 also show utilizing false negative data performs slightly better than false positives since false negative data might be predicted as positive relation and increase samples of the target relation to learn a more accurate decision boundary.

4.6 Case Study

We sample some examples of incorrectly labeled data which are regarded as unlabeled data during training. In Table 5, it can be seen that our discriminator recognizes both false positive and false negative instances. For example, though the fact (John Allison, *EmployedBy*, Opera) is absent in the KB due to the incompleteness of the KB, C2

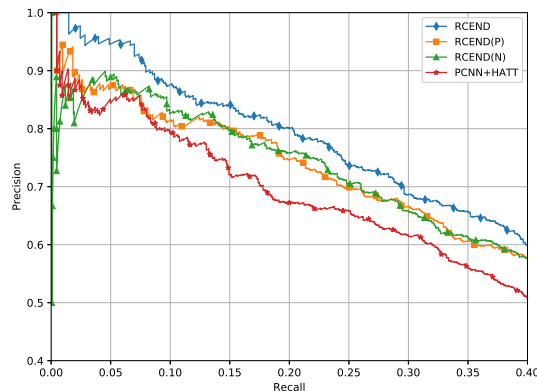


Figure 5: Precision-recall curves of our model with different settings.

expresses *EmployedBy* relation and provides evidence of target relation. Additionally, C4 is mislabeled as *BornIn* due to the relational fact (Bill Cosby, *BornIn*, Philadelphia), even though it mentions *LivedIn* relation. Further more, they are all predicted correctly by our relation classifier in the end which shows that our model indeed captures the valid information of noisy data and exploits them to enhance its ability.

5 Conclusion

In this paper, we proposed RCEND to fully exploit valid information of the noisy data in distant supervision relation classification. The instance discriminator is trained with reinforcement learning, which aims to recognize the instances mislabeled by distant supervision. We treat the correctly labeled instances as labeled data and incorrectly labeled ones as unlabeled data. Afterward, we adopt a semi-supervised learning method to learn a ro-

bust relation classifier to utilize the data. In this way, not only can our model reduce the side effect of noisy labels, but also adequately take advantage of valid information contained in noisy data. Experiments demonstrate that our model outperforms state-of-the-art baselines.

Acknowledgments

We would like to express gratitude to Robert Ridley and the anonymous reviewers for their valuable feedback on the paper. This work is supported by the National Natural Science Foundation of China (No. 61672277, U1836221), the Jiangsu Provincial Research Foundation for Basic Research (No. BK20170074).

References

- Dzmitry Bahdanau, Philemon Brakel, Kelvin Xu, Anirudh Goyal, Ryan Lowe, Joelle Pineau, Aaron C. Courville, and Yoshua Bengio. 2016. [An actor-critic algorithm for sequence prediction](#). *Computing Research Repository*, arXiv:1607.07086.
- Razvan Bunescu and Raymond Mooney. 2005. [A shortest path dependency kernel for relation extraction](#). In *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*.
- Jun Feng, Minlie Huang, Li Zhao, and Xiaoyan Yang, Yang amd Zhu. 2018. [Reinforcement learning for relation classification from noisy data](#).
- Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. [Incorporating non-local information into information extraction systems by gibbs sampling](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 363–370. Association for Computational Linguistics.
- Xu Han, Zhiyuan Liu, and Maosong Sun. 2018a. [Neural knowledge acquisition via mutual attention between knowledge graph and text](#). In *Proceedings of AAAI Conference on Artificial Intelligence*, pages 4832–4839.
- Xu Han, Pengfei Yu, Zhiyuan Liu, Maosong Sun, and Peng Li. 2018b. [Hierarchical relation extraction with coarse-to-fine grained attention](#). In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 2236–2245. Association for Computational Linguistics.
- Raphael Hoffmann, Congle Zhang, Xiao Ling, Luke Zettlemoyer, and Daniel S. Weld. 2011. [Knowledge-based weak supervision for information extraction of overlapping relations](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 541–550. Association for Computational Linguistics.
- Diederik P. Kingma and Jimmy Ba. 2014. [Adam: A method for stochastic optimization](#). *Computing Research Repository*, arXiv:1412.6980.
- Diederik P. Kingma, Shakir Mohamed, Danilo Jimenez Rezende, and Max Welling. 2014. [Semi-supervised learning with deep generative models](#). In *Advances in Neural Information Processing Systems*, pages 3581–3589.
- Yankai Lin, Shiqi Shen, Zhiyuan Liu, Huanbo Luan, and Maosong Sun. 2016. [Neural relation extraction with selective attention over instances](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 2124–2133. Association for Computational Linguistics.
- Tianyu Liu, Kexiang Wang, Baobao Chang, and Zhi-fang Sui. 2017. [A soft-label method for noise-tolerant distantly supervised relation extraction](#). In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1790–1795. Association for Computational Linguistics.
- Bingfeng Luo, Yansong Feng, Zheng Wang, Zhanxing Zhu, Songfang Huang, Rui Yan, and Dongyan Zhao. 2017. [Learning with noise: Enhance distantly supervised relation extraction with dynamic transition matrix](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 430–439. Association for Computational Linguistics.
- Mike Mintz, Steven Bills, Rion Snow, and Daniel Jurafsky. 2009. [Distant supervision for relation extraction without labeled data](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 1003–1011. Association for Computational Linguistics.
- Pengda Qin, Weiran XU, and William Yang Wang. 2018a. [Dsgan: Generative adversarial training for distant supervision relation extraction](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 496–505. Association for Computational Linguistics.
- Pengda Qin, Weiran XU, and William Yang Wang. 2018b. [Robust distant supervision relation extraction via deep reinforcement learning](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 2137–2147. Association for Computational Linguistics.
- Sebastian Riedel, Limin Yao, and Andrew McCallum. 2010. [Modeling relations and their mentions without labeled text](#). In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pages 148–163.

- David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Vedavyas Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy P. Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. 2016. [Mastering the game of go with deep neural networks and tree search](#). *Nature*, 529(7587):484–489.
- Mihai Surdeanu, Julie Tibshirani, Ramesh Nallapati, and Christopher D. Manning. 2012. [Multi-instance multi-label learning for relation extraction](#). In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 455–465. Association for Computational Linguistics.
- Richard S. Sutton, David A. McAllester, Satinder P. Singh, and Yishay Mansour. 1999. [Policy gradient methods for reinforcement learning with function approximation](#). In *Advances in Neural Information Processing Systems*, pages 1057–1063.
- Ronald J. Williams. 1992. [Simple statistical gradient-following algorithms for connectionist reinforcement learning](#). *Machine Learning*, 8:229–256.
- Kun Xu, Siva Reddy, Yansong Feng, Songfang Huang, and Dongyan Zhao. 2016. [Question answering on freebase via relation extraction and textual evidence](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*.
- Wei Xu, Raphael Hoffmann, Le Zhao, and Ralph Grishman. 2013. [Filling knowledge base gaps for distant supervision of relation extraction](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 665–670. Association for Computational Linguistics.
- Weidi Xu, Haoze Sun, Chao Deng, and Ying Tan. 2017. [Variational autoencoder for semi-supervised text classification](#). In *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pages 3358–3364.
- Dmitry Zelenko, Chinatsu Aone, and Anthony Richardella. 2002. [Kernel methods for relation extraction](#). In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.
- Daojian Zeng, Kang Liu, Yubo Chen, and Jun Zhao. 2015. [Distant supervision for relation extraction via piecewise convolutional neural networks](#). In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 1753–1762. Association for Computational Linguistics.
- Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. 2014. [Relation classification via convolutional deep neural network](#). In *Proceedings of International Conference on Computational Linguistics*, pages 2335–2344.
- GuoDong Zhou, Jian Su, Jie Zhang, and Min Zhang. 2005. [Exploring various knowledge in relation extraction](#). In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*, pages 427–434. Association for Computational Linguistics.