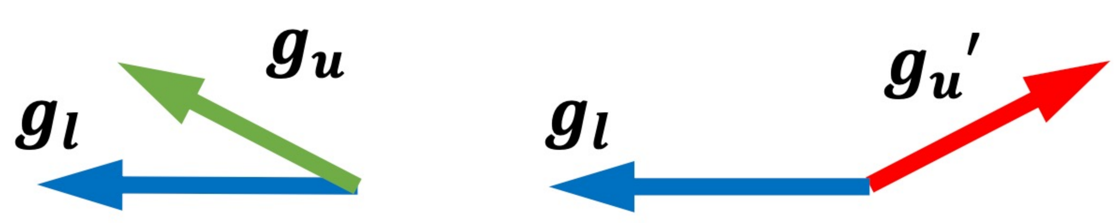
Gradient Imitation Reinforcement Learning for Low Resource Relation Extraction

Xuming Hu[1], Chenwei Zhang[2], Yawen Yang[1], Xiaohe Li[1], Li Lin[1]
Lijie Wen[1] and Philip S. Yu[1,3]
hxm19@mails.tsinghua.edu.cn
1 Tsinghua University 2 Amazon 3 University of Illinois at Chicago

Motivation



A letter$_{head}$ was delivered to my office$_{tail}$...

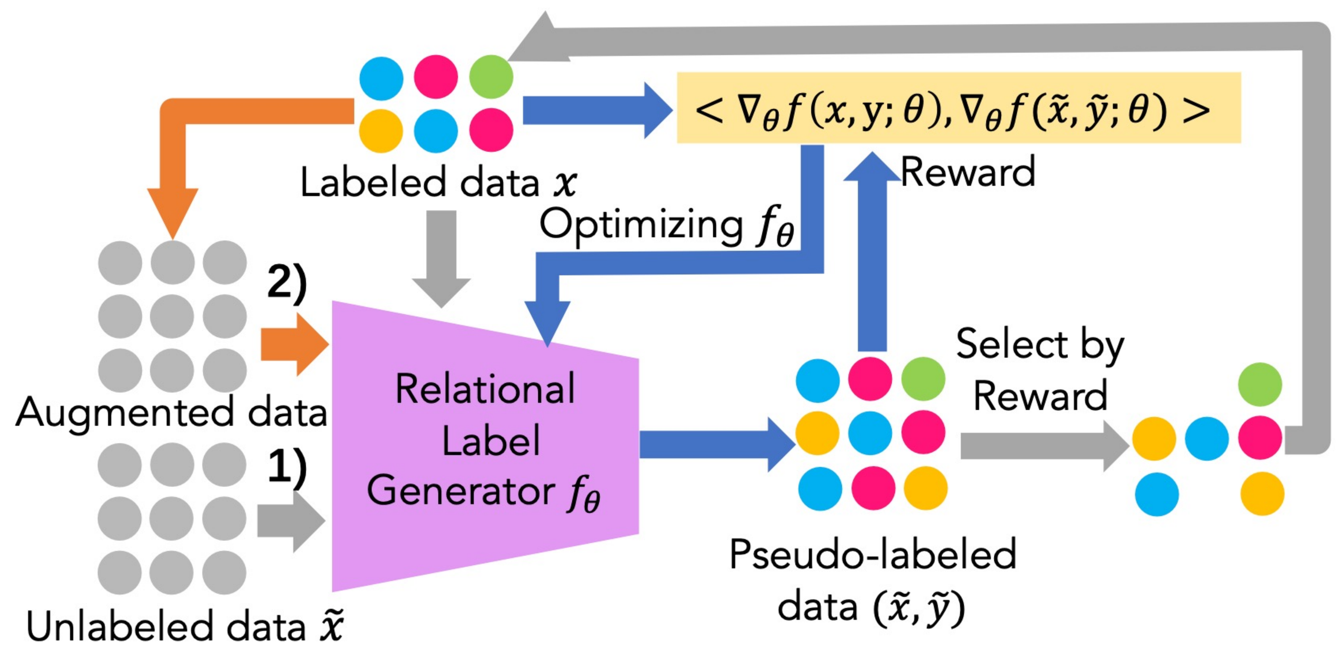
g_u : Entity-Destination ✓ Reward

g_u' : Entity-Origin ✗ Punishment

- Previous Methods: Directly used limited annotations during training.
- Our Motivation: Existing annotations as a guideline ➡ Unlabeled data to increases generalization ability.
- Our Method:

Design a reward ➡ Pseudo-labeled data mimics the gradient descent direction from limited labeled data ➡ *Reinforcement Learning*

Handle two major low resource scenarios



1. Limited labeled data and large amounts of unlabeled data are available. -> Gradient Imitation Reinforcement Learning module (GIRL)

Reinforcement Learning process:

- **State:** Updated labeled dataset D_l and standard gradient direction g_l at step t.
- **Policy:** RLG network f_θ.
- **Action:** Predict relational label on unlabeled data $\tilde{x}^{(t)}$ as pseudo-labeled data $(\tilde{x}^{(t)}, \tilde{y}^{(t)})$ at step t.
- **Reward:** Standard gradient descent direction on the all N labeled data.

$$g_l^{(n)}(\theta) = \nabla_\theta \mathcal{L}_l(x^{(n)}, y^{(n)}; \theta)$$

Expected gradient descent direction on the pseudo-labeled data.

$$g_p^{(t)}(\theta) = \nabla_\theta \mathcal{L}_p(\tilde{x}^{(t)}, \tilde{y}^{(t)}; \theta)$$

Cosine similarity between g_l and g_p for state $s^{(t)}$.

$$R^{(t)} = \frac{g_l(\theta)^T g_p(\theta)}{||g_l(\theta)||_2 ||g_p(\theta)||_2}$$

- **Update State:** For these positive reinforcement $R^{(t)}$>0.5:

$$D_l \leftarrow D_l \cup D_p$$

$$g_l \leftarrow \frac{1}{N+1}(N g_l + g_p)$$

- **Reinforcement learning loss:** We calculate the loss over a batch of pseudo-labeled samples.

$$\mathcal{L}(\theta) = \sum_{t=1}^{T} loss\left(f_\theta\left(\tilde{x}^{(t,E1,E2)}\right), \text{one\_hot}(\tilde{y}^{(t)})\right) * R^{(t)}$$

2. Only limited labeled data is available. -> Contextualized Data Augmentation module (CDA)

We sample spans of the sentence as [MASK] and finally fills the mask with tokens using BERT.

A letter was delivered to my office in this morning.

⬇ Sample spans as [MASK]

A letter was [MASK] [MASK] my office in this morning.

⬇ Fill the [MASK]

A letter was sent from my office in this morning.

Experiments

Datasets

Datasets	SemEval	TACRED
Relation mentions	7199/800/1864	75049/25763/18659
Relation	19	42
No_relation rate	17.4%	78.7%

Implementation

Datasets	SemEval	TACRED
Labeled set	5%/10%/30%	3%/10%/15%
Unlabeled set	50%	50%

Baselines

- Relation Encoders
  - LSTM (Hochreiter and Schmidhuber, 1997)
  - PCNN (Zeng et al., 2015)
  - PRNN (Zhang et al., 2017)
  - BERT (Devlin et al., 2019)
- Self-Training (Rosenberg et al., 2005)
- Mean-Teacher (Tarvainen and Valpola, 2017)
- DualRE (Lin et al., 2019)
- RE-Ensemble (Lin et al., 2019)
- MRefG (Li and Qian, 2020)
- MetaSRE (Hu et al., 2021)
- BERT w. gold labels

Does GIRL helps to improve pseudo label quality? Yes!

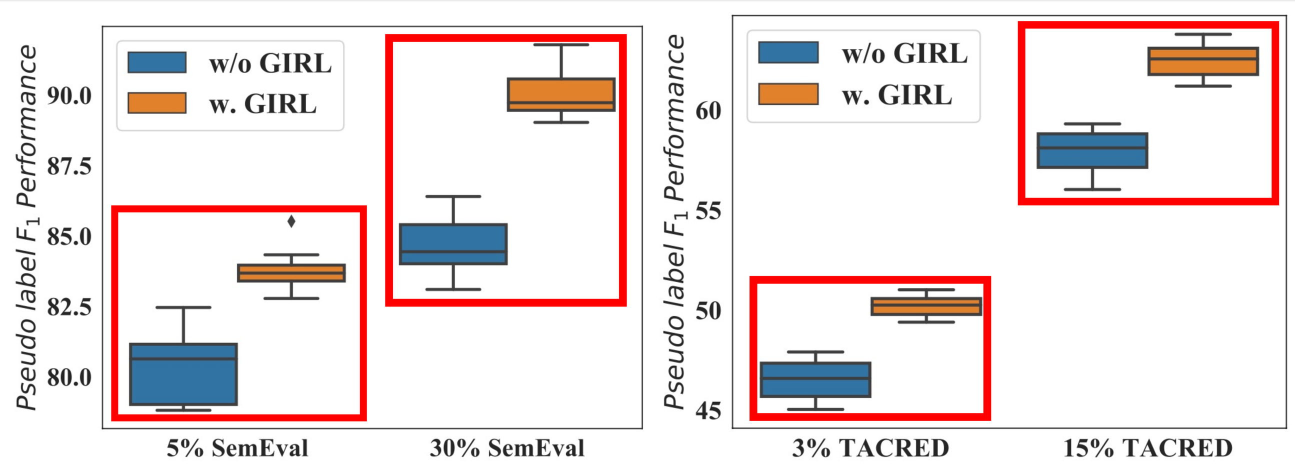


Figure 2: Pseudo label F1 (%) Performance with GIRL based on SemEval (left) and TACRED (right).

From Figure 2, we could find for the two datasets with different ratios of the labeled data, GIRL could undoubtedly improve the F1 performance of pseudo labels.

Does GIRL helps to guide the gradient descent direction? Yes!

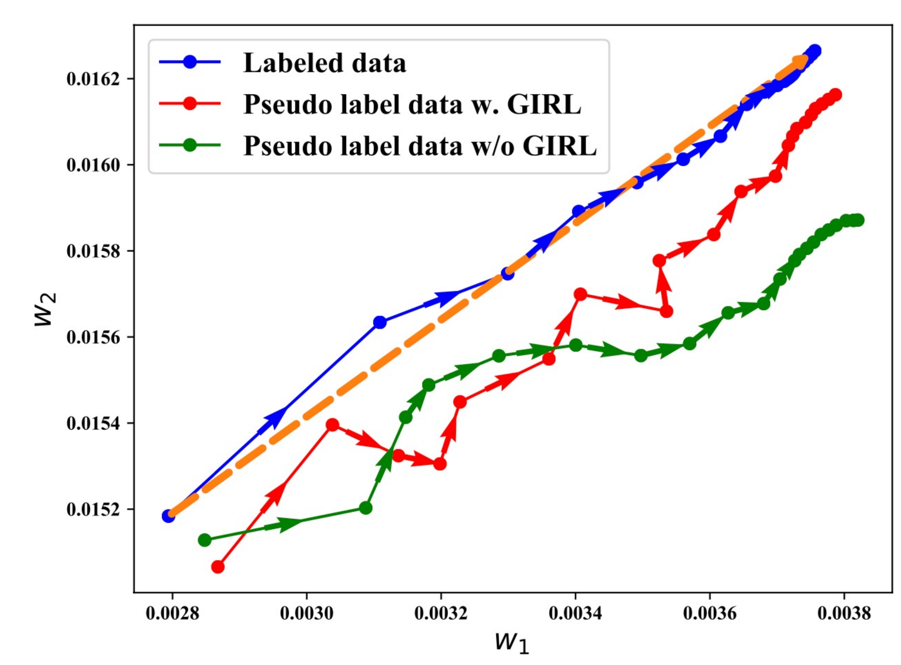


Figure 3: GradLRE gradient descent directions on labeled data and pseudo label data. The dotted line indicates the average gradient direction on labeled data.

In Figure 3, the optimization direction of pseudo label data fluctuates at the beginning, GIRL is gradually improving and ends up closer to the ideal local minima. When GIRL is not used, the error-prone pseudo labels obtained without instructive feedback gradually push the optimization away from the local minima

Two major low resource scenarios results.

	% Labeled Data	L	L + CDA	L + U
SemEval	5%	72.71	75.52	79.65
	10%	73.93	81.47	81.69
	30%	80.55	84.63	85.52
TACRED	3%	41.11	43.34	47.37
	10%	53.23	57.07	58.20
	15%	55.35	58.89	59.93

Table 3: F1 (%) of GradLRE with various percentages of labeled data under different LRE scenarios.

1)L+U: Limited labeled data + 50% unlabeled data.
2)L+CDA: Limited labeled data + CDA generate 50% unlabeled data.
3)L: Limited labeled data.

We present results in Table 3. Compared to L, L+CDA achieves an average 4.01% improvement in F1, indicating the effectiveness of contextualized data augmentation. We also observe that L+CDA obtain competitive performance when compared with L+U on SemEval.

Does CDA generate useful unlabeled data? Yes!

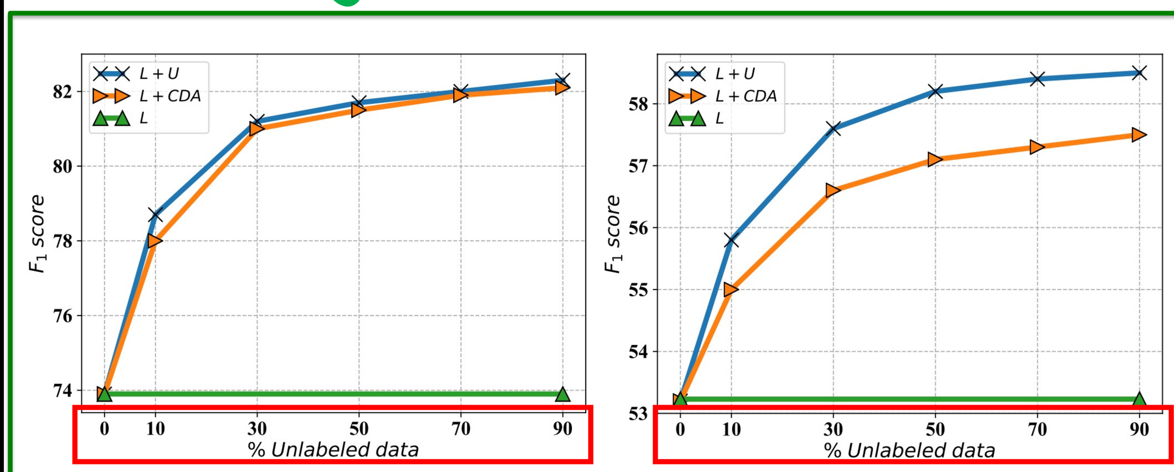


Figure 4: F1 (%) Performance with various unlabeled data and 10% labeled data on SemEval (left) and TACRED (right).

From Figure 4, L+CDA outperforms L consistently, with the ratio of unlabeled data increasing, L+CDA can get more discriminative data and obtain better performance: it can achieve almost the same performance as L+U on SemEval.

Case study

GIRL

My *brother* has entered my *room* without knocking.
Label: `Entity-Destination`
Prediction w/o GIRL: `Other`
Prediction w. GIRL: `Entity-Destination`

Ditto for his funny turn as a *man* who instigates the *kidnapping* of his own wife in ...
Label: `Cause-Effect`
Prediction w/o GIRL: `Other`
Prediction w. GIRL: `Cause-Effect`

Other -> Correct labels

CDA

Original: A *letter* was *delivered to* my *office* in ...
Label: `Entity-Destination`
Generated: A *letter* was *sent from* my *office* in ...
Pseudo label: `Entity-Origin`

Change the original relation

Original: The *editor* improved the *manuscript* with *his changes*.
Label: `Product-Producer`
Generated: The *editor* improved the *manuscript* with *some improvements*.
Pseudo label: `Product-Producer`

Maintain the original relation