Gradient Imitation Reinforcement Learning for Low Resource Relation Extraction

Xuming Hu[1], Chenwei Zhang[2], Yawen Yang[1], Xiaohe Li[1], Lin Li[1], Lijie Wen[1] ,Philip S. Yu[1,3]

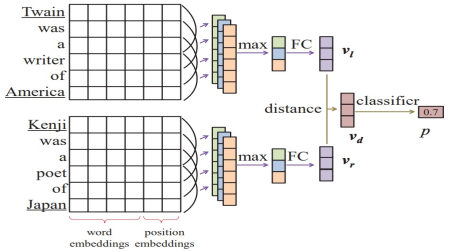
[1] Tsinghua University
[2] Amazon
[3] University of Illinois at Chicago

Relation Extraction

Labor-intensive

Sentence		Relation
Derek Bell was born in *Belfast*. *Donald Trump* was born in *America*. ……	→	Born In
Thomson is based in *Toronto*. *Beijing* is located in *China*. ……	→	Located in

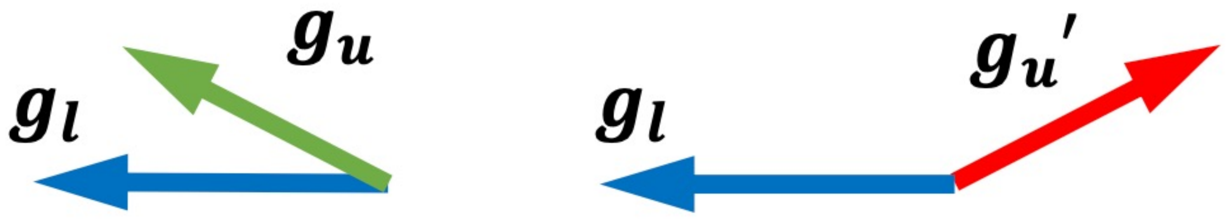
Relation Encoder + Deep Classification Model



(Stanovsky et al., 2018; Saha et al., 2018; Yu et al., 2017)

How to improve the model performance for LRE?

- Previous Methods: Directly used limited annotations during training.
- Shortage: The trained models inevitably possesses selection bias.
- Motivation: How to use existing annotations as a guideline and leverage unlabeled data to increases generalization ability?



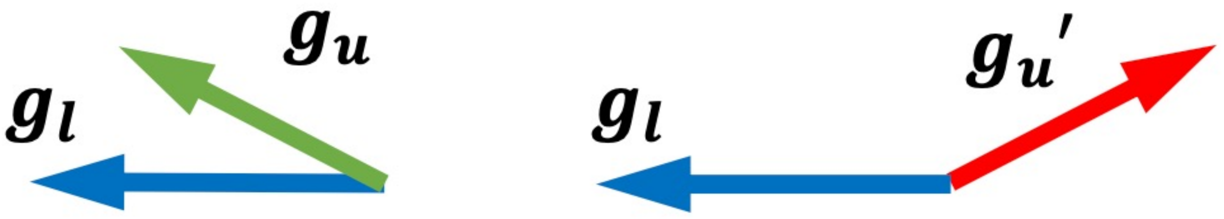
A letter$_{head}$ was delivered to my office$_{tail}$...

g_u : Entity-Destination ✓ Positive

g_u' : Entity-Origin ✗ Negative

How to improve the model performance for LRE?

Design a reward → Explicit feedback → *Reinforcement learning*



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

Reward g_u : **Entity-Destination** Positive

Punishment g_u' : **Entity-Origin** Negative

Framework (GradLRE)

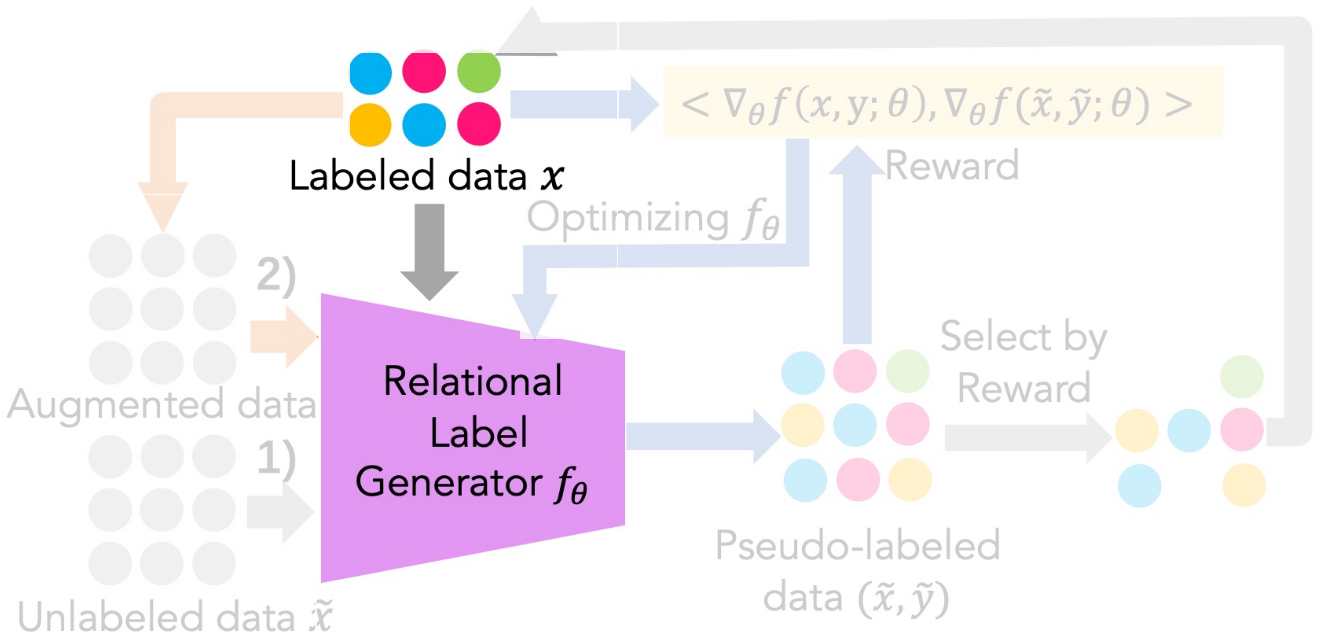
1) Limited labeled data and large amounts of unlabeled data are available

- Relation Label Generator (RLG)
- Gradient Imitation Reinforcement Learning (GIRL)

2) Only limited labeled data is available

- Contextualized Data Augmentation (CDA)

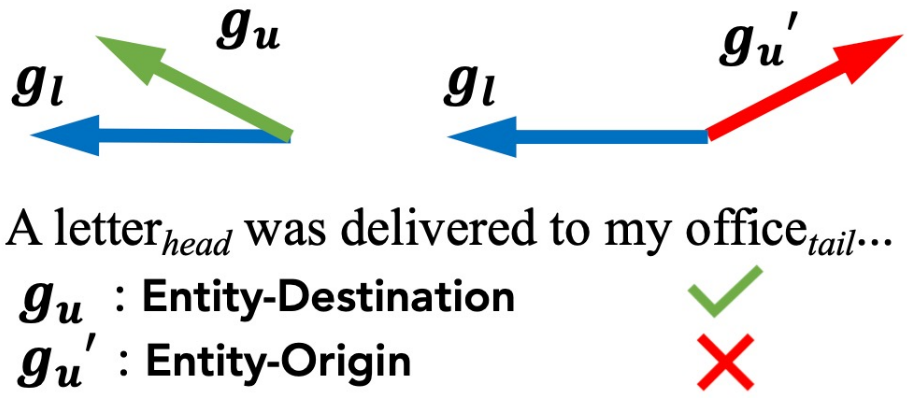
Relation Label Generator (RLG)



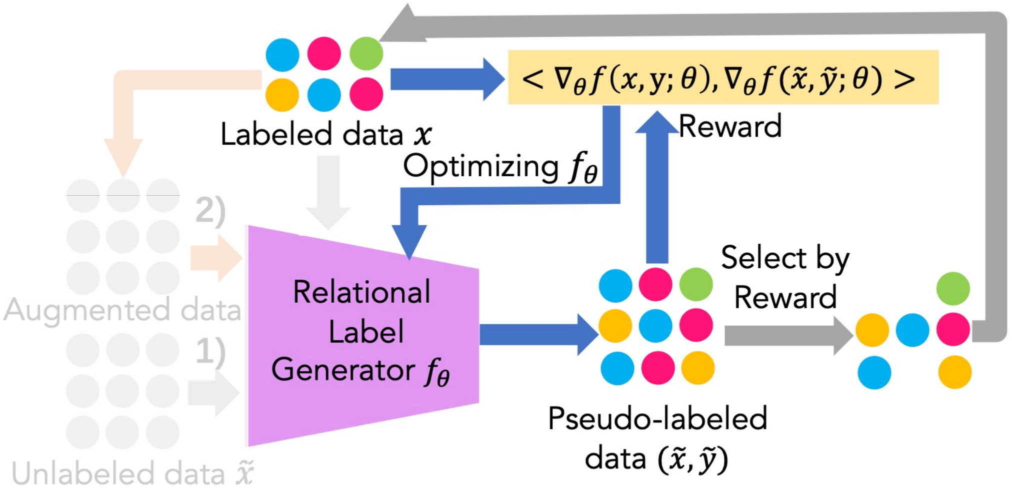
- Mark two eneities with four reserved tokens $[E1]$, $[/E1]$, $[E2]$, $[/E2]$:
  $A\ [E1]\ letter\ [/E1]\ was\ delivered\ to\ my\ [E2]\ office[/E2]\ \ldots$
- Get the relation representation of two entities corresponding to $[E1]$,$[E2]$ from BERT.
  $$\mathbf{h} = [\mathbf{h}_{[E1]}, \mathbf{h}_{[E2]}]$$
- Classifiy these representations into specific relations with a fully connected network $f_\theta(x, E1, E2)$.

Gradient Imitation Reinforcement Learning (GIRL)

- Define Standard gradient descending:
  Partial derivatives on the labeled data $\nabla_\theta f(x, y; \theta)$
- Assume: When pseudo-labeled data are correctly labeled, partial derivatives on the pseudo-labeled data would be highly similar to standard gradient descending.



Gradient Imitation Reinforcement Learning (GIRL)



- 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.

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

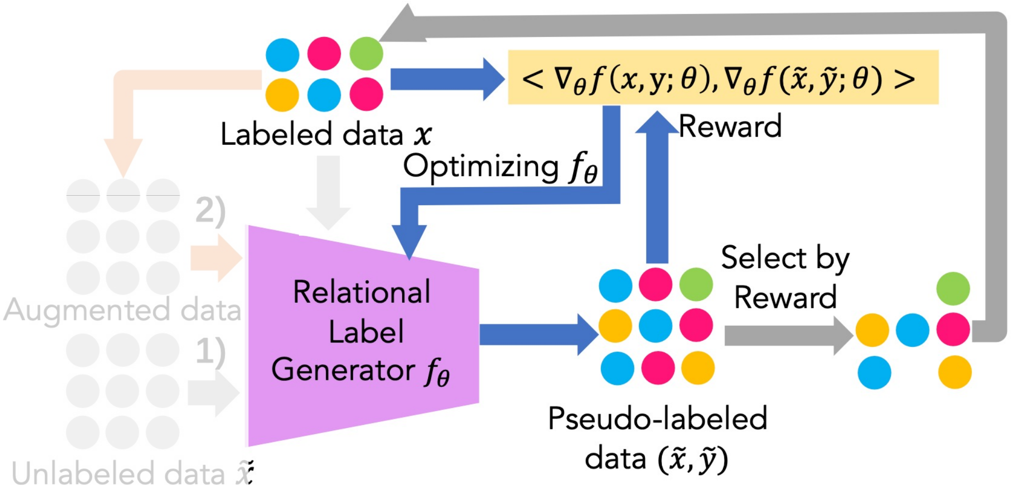
  Expected gradient descent direction on the pseudo-labeled data.

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

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

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

Gradient Imitation Reinforcement Learning (GIRL)



- ## 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

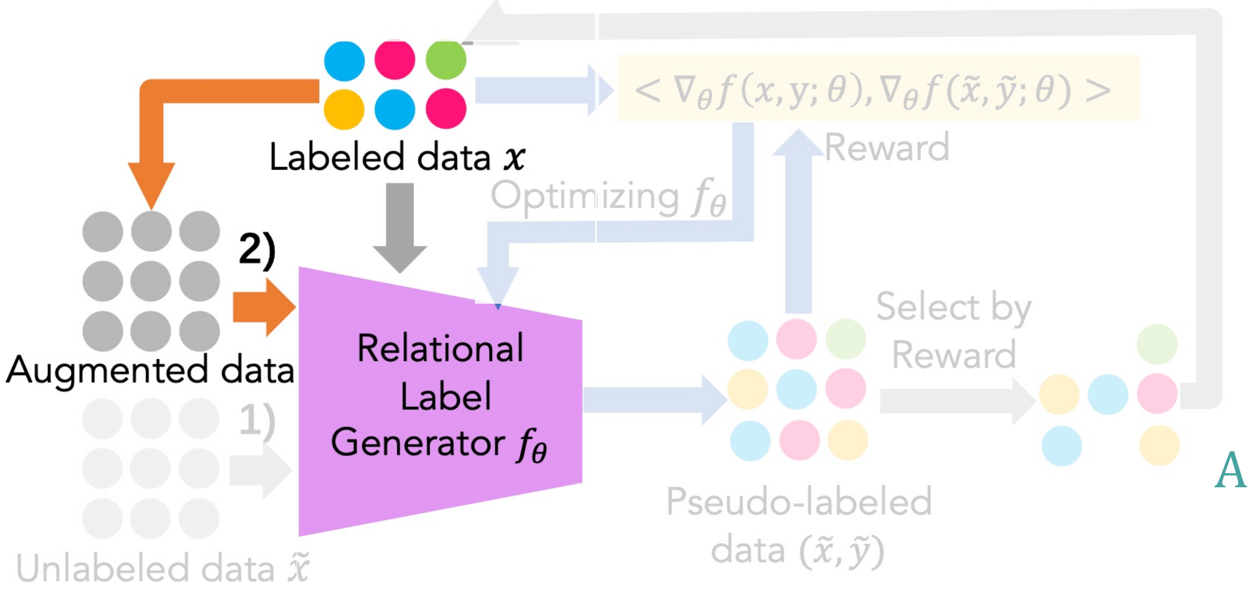
We calculate the loss over a batch of pseudo-labeled samples.

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

Contextualized Data Augmentation

Generate unlabeled data

Only limited labeled data is av



- CDA samples 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%

Implementations

Datasets	SemEval	TACRED
Labeled set Unlabeled set	5%/10%/30% 50%	3%/10%/15% 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 pseud… …lity?

Yes!

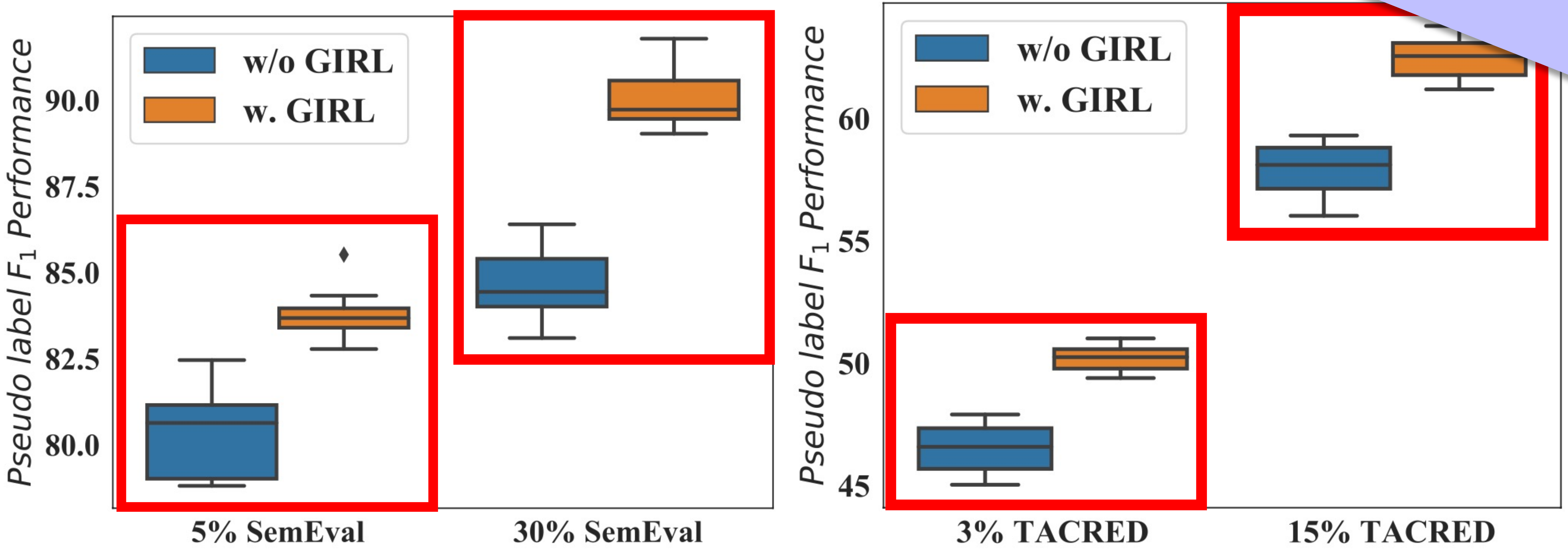


Figure 2: Pseudo label F1 (%) Performance with GIRL based on SemEval (left) and TACRED (right).

Does GIRL helps to guide the gradient des n ?

Yes!

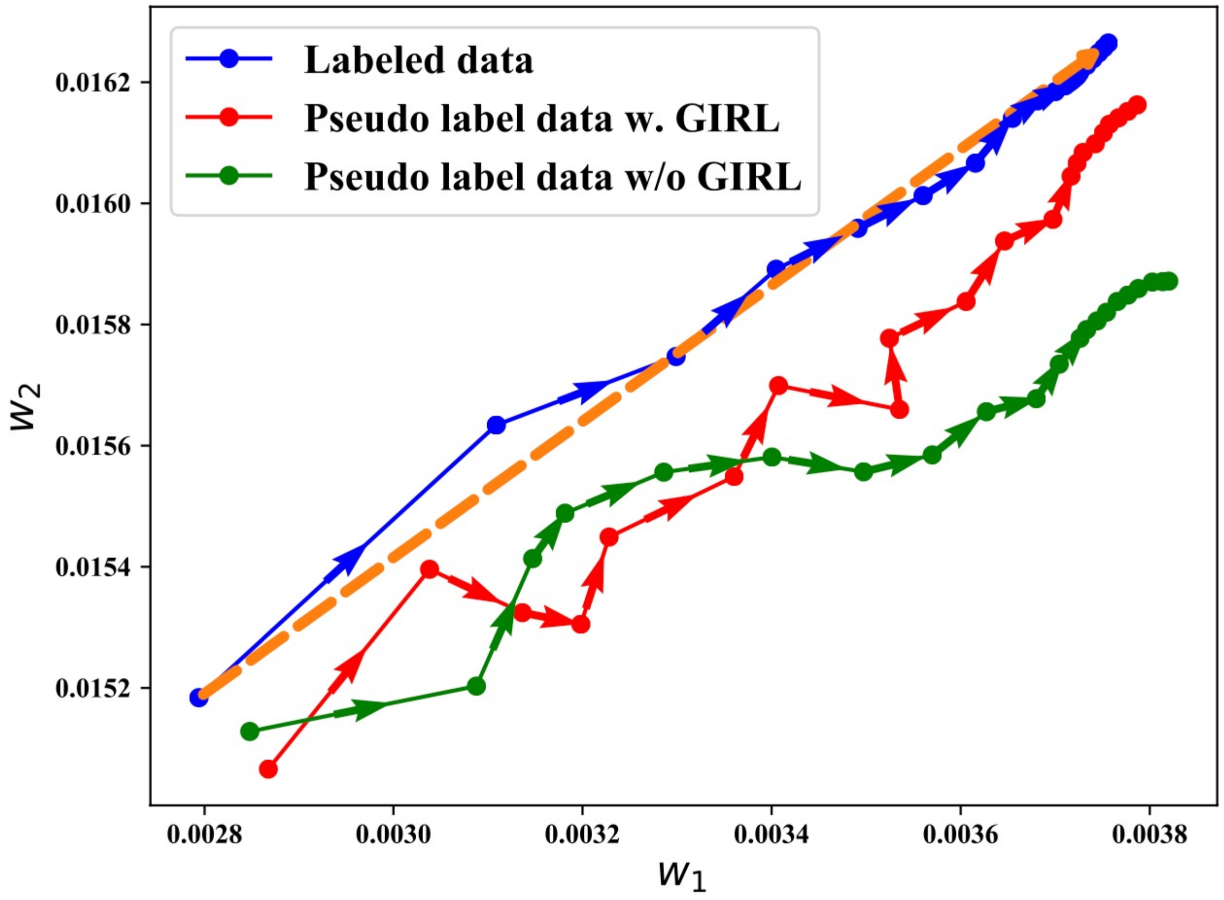


Figure 3: GradLRE gradient descent directions on labeled data and pseudo label data. The dotted line indicates the average gradient direction on labeled data.

Case study using GIRL

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

The *disc* in a disc *music box* plays this function, with pins perpendicular to the plane surface...
Label: **Content-Container**
Prediction w/o GIRL: **Component-Whole**
Prediction w. GIRL: **Content-Container**

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**

Table 2: Predictions with/without GIRL on SemEval, where *red* and *blue* represent head and tail entities respectively.

Other -> Correct labels

Handling two major low resource 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.

	% 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.

Does CDA generate useful unlabel

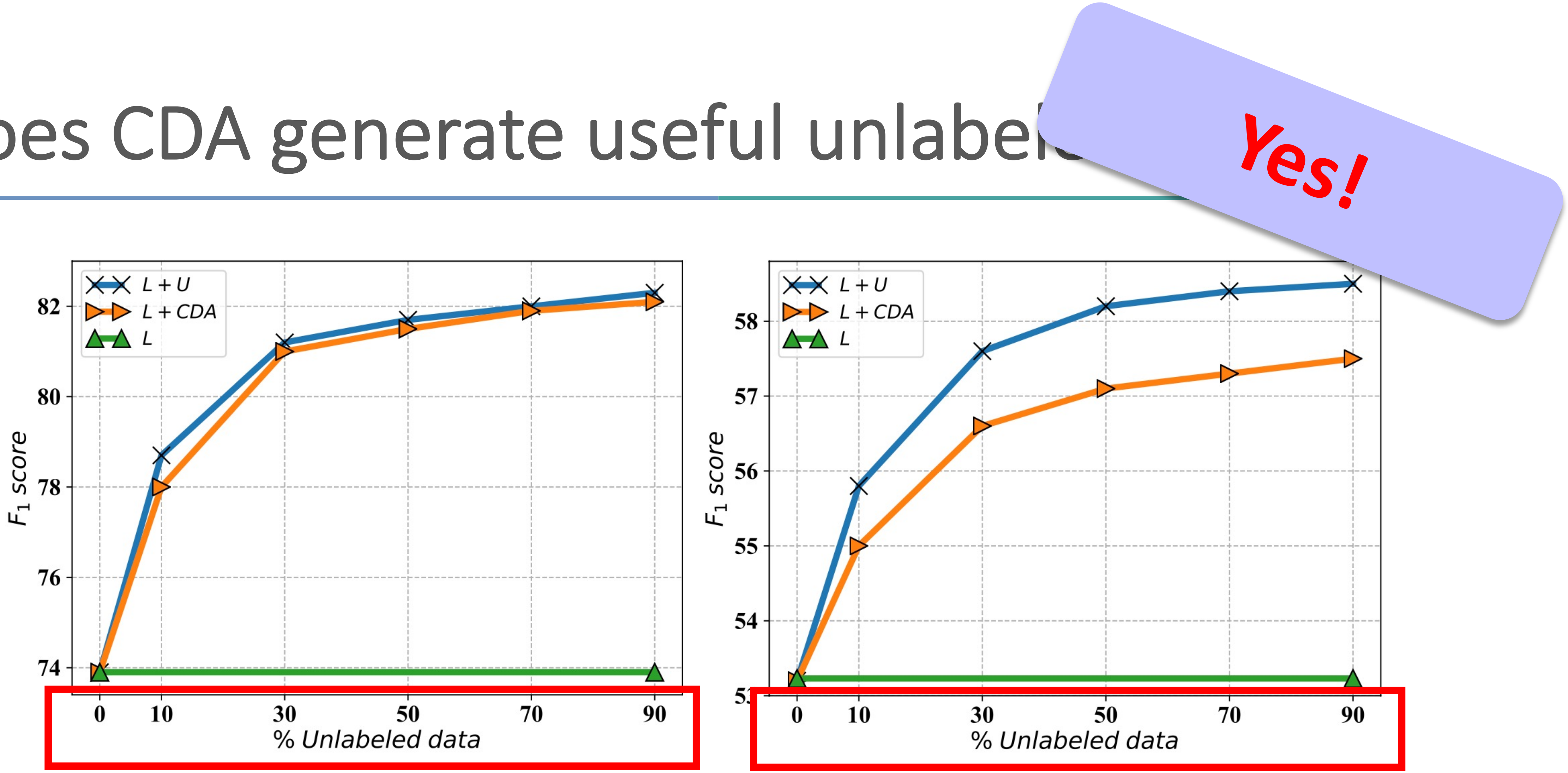


Figure 4: F1 (%) Performance with various unlabeled data and 10% labeled data on SemEval (left) and TACRED (right).

Case study using CDA

Original: A *letter* was *delivered to* my *office* in ...	
Label: **Entity-Destination**	Maintain the original relation
Generated: A *letter* was *sent from* my *office* in ...	
Pseudo label: **Entity-Origin**	
Original: The *editor* improved the *manuscript* with *his changes*.	
Label: **Product-Producer**	Change the original relation
Generated: The *editor* improved the *manuscript* with *some improvements*.	
Pseudo label: **Product-Producer**	
Original: The *suspect dumped* the *dead body* into a local *reservoir*.	
Label: **Entity-Destination**	
Generated: The *dam bulids* the *human body* into a local *reservoir*.	
Pseudo label: **Other**	

Table 4: CDA on labeled data to obtain generated data, where *red* and *blue* represent head and tail entities respectively, *cyan* represents the replaced words.

Conclusion

- Our model encourages pseudo-labeled data to imitate the gradient optimization direction in labeled data to improve the pseudo label quality.
- Contextualized data augmentation is proposed to handle the extremely low resource Relation Extraction situation where no unlabeled data is available.
- Experiments on two public datasets show effectiveness of GradLRE and augmented data over competitive baselines.

THANK YOU!

Code + Data are Available at:
http://github.com/THU-BPM/GradLRE