Gradient Imitation Reinforcement Learning for Low Resource Relation Extraction

Xuming Hu\textsuperscript{1}, Chenwei Zhang\textsuperscript{2}, Yawen Yang\textsuperscript{1}, Xiaohe Li\textsuperscript{1}, Lin Li\textsuperscript{1}, Lijie Wen\textsuperscript{1}, Philip S. Yu\textsuperscript{1,3}

\textsuperscript{1} Tsinghua University
\textsuperscript{2} Amazon
\textsuperscript{3} University of Illinois at Chicago
Relation Extraction

Sentence

Derek Bell was born in Belfast.

Donald Trump was born in America.

Thomson is based in Toronto.

Beijing is located in China.

......

Relation

Born In

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 possess selection bias.
- Motivation: How to use existing annotations as a guideline and leverage unlabeled data to increase generalization ability?

A letter was delivered to my office...

- $g_u$: Entity-Destination
- $g_u'$: Entity-Origin

A letter was not delivered to my office...

- $g_l$: Entity-Origin
- $g_u'$: Entity-Destination

A letter was delivered to my office...

- $g_l$: Entity-Destination
- $g_u':$ Entity-Origin

Positive

Negative
How to improve the model performance for LRE?

Design a reward → Explicit feedback → Reinforcement learning

A letter $g_l$ was delivered to my office $g_u$...

Reward $g_u':$ Entity-Origin
Punishment $g_u':$ Entity-Destination

Positive
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 entities with four reserved tokens \([E1], [/E1], [E2], [/E2]\):
  A \([E1]\) letter \([/E1]\) was delivered to my \([E2]\) office \([/E2]\) ...
- Get the relation representation of two entities corresponding to \([E1],[E2]\) from BERT.
  \(h = [h_{[E1]}, h_{[E2]}]\)
- Classify 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 $f_\theta$.

• 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 $R(t)>0.5$:
  
  $$D_l \leftarrow D_l \cup D_p$$
  $$g_l \leftarrow \frac{1}{N+1} (Ng_l + g_p)$$

- **Reinforcement Learning loss**

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

  $$L(\theta) = \sum_{t=1}^{T} \text{loss}(f_\theta(\tilde{x}^{(t,F1,F2)}), \text{one_hot}(\tilde{y}^{(t)})) * R(t)$$
Contextualized Data Augmentation (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

<table>
<thead>
<tr>
<th>Datasets</th>
<th>SemEval</th>
<th>TACRED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relation mentions</td>
<td>7199/800/1864</td>
<td>75049/25763/18659</td>
</tr>
<tr>
<td>Relation</td>
<td>19</td>
<td>42</td>
</tr>
<tr>
<td>No_relation rate</td>
<td>17.4%</td>
<td>78.7%</td>
</tr>
</tbody>
</table>

### 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**

### Implementations

<table>
<thead>
<tr>
<th>Datasets</th>
<th>SemEval</th>
<th>TACRED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Labeled set</td>
<td>5%/10%/30%</td>
<td>3%/10%/15%</td>
</tr>
<tr>
<td>Unlabeled set</td>
<td>50%</td>
<td>50%</td>
</tr>
</tbody>
</table>

---

11/7/21 Gradient Imitation Reinforcement Learning for Low Resource Relation Extraction
Figure 2: Pseudo label F1 (%) Performance with GIRL based on SemEval (left) and TACRED (right).
Does GIRL help 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.
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.
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.

<table>
<thead>
<tr>
<th>% Labeled Data</th>
<th>L</th>
<th>L + CDA</th>
<th>L + U</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>72.71</td>
<td>75.52</td>
<td>79.65</td>
</tr>
<tr>
<td>SemEval</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>73.93</td>
<td>81.47</td>
<td>81.69</td>
</tr>
<tr>
<td>30%</td>
<td>80.55</td>
<td>84.63</td>
<td>85.52</td>
</tr>
<tr>
<td>3%</td>
<td>41.11</td>
<td>43.34</td>
<td>47.37</td>
</tr>
<tr>
<td>TACRED</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10%</td>
<td>53.23</td>
<td>57.07</td>
<td>58.20</td>
</tr>
<tr>
<td>15%</td>
<td>55.35</td>
<td>58.89</td>
<td>59.93</td>
</tr>
</tbody>
</table>

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

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

Original: The suspect dumped the dead body into a local reservoir.
Label: Entity-Destination
Generated: The dam builds 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