Heterogeneous Supervision for Relation Extraction: A Representation Learning Approach

Liyuan Liu*, Xiang Ren*, Qi Zhu*, Huan Gui*, Shi Zhi*, Heng Ji and Jiawei Han*  

*University of Illinois at Urbana-Champaign, Urbana, IL, USA  
◊Computer Science Department, Rensselaer Polytechnic Institute, USA
Relation Extraction

• Goal: acquire structured knowledge from unstructured text

“Hussein was born in Amman on 14 November 1935.”
Relation Extraction

• Formal Definition:
  • Sentence-level relation extraction:
    • Classify a relation mention into a set of relation types of interest or Not-Target-Type (None)
Relation Extraction

• Formal Definition:
  • Sentence-level relation extraction:
    • Classify a relation mention into a set of relation types of interest or Not-Target-Type (None)

Hussein was born in Amman on 14 November 1935.

entity pair

Hussein Amman

sentence / context

multi-class classification

Born-in
President-of
Died-in
Parents-of
...
None
Related Work

• Supervised Learning:
  • Multi-class classification
Related Work

• Supervised Learning:
  • Multi-class classification

Dataset with human annotation is the bottleneck

Limited, might even not existed for many domains
Hard to get, and costly
Slow, and sometimes outdated
......
Related Work

• Bootstrap learning:
  • Start with a set of seed patterns / annotations, iteratively generate more
  • Suffers from semantic shift

Mintz et al. “Distant supervision for relation extraction without labeled data”, ACL 2009
Related Work

• Distant Supervision:
  • Automatically generate annotations by Knowledge Base

Mintz et al. “Distant supervision for relation extraction without labeled data”, ACL 2009
Related Work

• Distant Supervision:

  • Automatically generate annotations by Knowledge Base
    • ("Obama", "USA", Obama was born in Honolulu, Hawaii, USA as he has always said)
      • Born-in (correct)
      • President-of (wrong).
Related Work

• Distant Supervision:
  • Automatically generate annotations by Knowledge Base

Distant supervision only encodes KB, while we have more than KB
Heterogeneous Supervision

• Provide a general framework to encode knowledge for supervision:
  • Knowledge Base, domain-specific patterns, ......

• Labelling functions:

\[
\begin{align*}
\lambda_1 & \quad \text{return } \textit{born\_in} \text{ for } \langle e_1, e_2, s \rangle \text{ if } \text{BornIn}(e_1, e_2) \text{ in KB} \\
\lambda_2 & \quad \text{return } \textit{died\_in} \text{ for } \langle e_1, e_2, s \rangle \text{ if } \text{DiedIn}(e_1, e_2) \text{ in KB} \\
\lambda_3 & \quad \text{return } \textit{born\_in} \text{ for } \langle e_1, e_2, s \rangle \text{ if match}(' * \text{ born in } *', s) \\
\lambda_4 & \quad \text{return } \textit{died\_in} \text{ for } \langle e_1, e_2, s \rangle \text{ if match}(' * \text{ killed in } *', s)
\end{align*}
\]

Heterogeneous Supervision & Distant Supervision:

• Heterogeneous Supervision is an extension of Distant Supervision:
  • Both encode external information and provide supervision,
  • Heterogeneous Supervision can encode more.

<table>
<thead>
<tr>
<th>Information type</th>
<th>KBP</th>
<th>NYT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of Relation Types</td>
<td># of Relation Mentions</td>
</tr>
<tr>
<td>Knowledge Base</td>
<td>7</td>
<td>133955</td>
</tr>
<tr>
<td>Domain-specific Patterns</td>
<td>13</td>
<td>225977</td>
</tr>
</tbody>
</table>

Table1. Statistic of Heterogeneous Supervision
Heterogeneous Supervision & Distant Supervision:

• Heterogeneous Supervision is an extension of Distant Supervision:
  • Both encode external information and provide supervision,
  • Heterogeneous Supervision can encode more.

<table>
<thead>
<tr>
<th>Information type</th>
<th>KBP</th>
<th>NYT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># of Relation Types</td>
<td># of Relation Mentions</td>
</tr>
<tr>
<td>Knowledge Base</td>
<td>7</td>
<td>133955</td>
</tr>
<tr>
<td>Domain-specific Patterns</td>
<td>13</td>
<td>225977</td>
</tr>
</tbody>
</table>

Table1. Statistic of Heterogeneous Supervision
Challenges

• Relation Extraction
• Resolve Conflicts among Heterogeneous Supervision
ReHession

• Heterogeneous Supervision

• Our Solution: A Representation Learning Approach
  • Relation Mention Representation
  • True Label Discovery component
  • Relation Extraction component

• Experiments
Conflicts among Heterogeneous Supervision

• Most simple way: majority voting

\[ D \]

\[ C_1 \] Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader Jesse James \( (e_1) \) in Missouri \( (e_2) \).

\[ C_2 \] Gofraid \( (e_1) \) died in 989, said to be killed in Dal Riata \( (e_2) \).

\[ C_3 \] Hussein \( (e_1) \) was born in Amman \( (e_2) \) on 14 November 1935.

\[ \Lambda \]

\[ \lambda_1 \] return \text{born\_in} for \( <e_1,e_2,s> \) if \text{BornIn}(e_1,e_2) in KB

\[ \lambda_2 \] return \text{died\_in} for \( <e_1,e_2,s> \) if \text{DiedIn}(e_1,e_2) in KB

\[ \lambda_3 \] return \text{born\_in} for \( <e_1,e_2,s> \) if match(" * born in * ", s)

\[ \lambda_4 \] return \text{died\_in} for \( <e_1,e_2,s> \) if match(" * killed in * ", s)
Conflicts among Heterogeneous Supervision

• How to resolve conflicts among Heterogeneous Supervision?
• Works for C3 and C2, but not work for C1
Conflicts among Heterogeneous Supervision

• For more complicated models, several principles have been proposed:
  • Truth Discovery:
    • Some sources (labeling functions) would be more reliable than others
    • Refer the reliability of different sources and the true label at the same time
    • Source Consistency Assumption: a source is likely to provide true information with the same probability for all instances.

Conflicts among Heterogeneous Supervision

• For more complicated models, several principles have been proposed:
  • Truth Discovery:
    • May not fit our scenario very well

Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader **Jesse James** (e₁) in **Missouri** (e₂).

**Gofraid** (e₁) died in 989, said to be killed in **Dal Riata** (e₂).

**Hussein** (e₁) was born in **Amman** (e₂) on 14 November 1935.
Conflicts among Heterogeneous Supervision

• For more complicated models, several principles have been proposed:
  • Truth Discovery:
    • These models are context-agnostic, while context is important for Relation Extraction
Conflicts among Heterogeneous Supervision

• For more complicated models, several principles have been proposed:
  • Truth Discovery:
  • Distant Supervision:
    • Partial-label association has been proposed to resolve conflicts among Distant Supervision, and proved to be effective.

\[ l(z, O_z) = \max\{0, 1 - \left[ \max_{r \in O_z} \phi(z, r) - \max_{r' \notin O_z} \phi(z, r') \right] \} \]

Most likely positive relation type

Most likely negative relation type

Conflicts among Heterogeneous Supervision

• For more complicated models, several principles have been proposed:
  • Truth Discovery:
  • Distant Supervision:
    • Partial-label association has been proposed to resolve conflicts among Distant Supervision, and proved to be effective.
    \[ l(z, O_z) = \max\{0, 1 - \left[ \max_{r \in O_z} \phi(z, r) - \max_{r' \notin O_z} \phi(z, r') \right] \} \]
  • For Distant Supervision, all annotations come from Knowledge Base.
  • For Heterogeneous Supervision, annotations are from different sources, and some could be more reliable than others.
Conflicts among Heterogeneous Supervision

• To fit our problem, we introduce context awareness to truth discovery, and modified the assumption:

  • A source is likely to provide true information with the same probability for instances \textit{with similar context}. 

23
Heterogeneous Supervision

• To fit our assumption, we add one constraint to labeling functions:
  • each labeling function can annotate only one relation type based on one source of information

• Reasons:
  • Different information sources often have different reliabilities

  • Some sources annotate different relation types without consistency
    • KB-based labeling function may have higher recall on ‘president-of’ than ‘born-in’
Heterogeneous Supervision

• To fit our assumption, we add one constraint to labeling functions:
  • each labeling function can annotate only one relation type based on one source of information

return $r$ for $<e_1, e_2, s>$ if $r(e_1, e_2)$ in KB
Heterogeneous Supervision

- To fit our assumption, we add one constraint to labeling functions:
  - each labeling function can annotate only one relation type based on one source of information

\[
\begin{align*}
\lambda_1 & \text{return } \textit{born_in} \text{ for } <e_1, e_2, s> \text{ if } \text{BornIn}(e_1, e_2) \text{ in KB} \\
\lambda_2 & \text{return } \textit{died_in} \text{ for } <e_1, e_2, s> \text{ if } \text{DiedIn}(e_1, e_2) \text{ in KB} \\
\lambda_3 & \text{return } \textit{born_in} \text{ for } <e_1, e_2, s> \text{ if } \text{match}(\text{"* born in *"}, s) \\
\lambda_4 & \text{return } \textit{died_in} \text{ for } <e_1, e_2, s> \text{ if } \text{match}(\text{"* killed in *"}, s)
\end{align*}
\]
ReHession

• Heterogeneous Supervision

• Our Solution: A Representation Learning Approach
  • Relation Mention Representation
  • True Label Discovery component
  • Relation Extraction component

• Experiments
Heterogeneous Supervision for Relation Extraction

- **Relation Extraction:**
  - Matching context with proper relation type

- **Heterogeneous Supervision:**
  - Refer true labels in a context-aware manner
Heterogeneous Supervision for Relation Extraction

• Relation Extraction:
  • Matching context with proper relation type

• Heterogeneous Supervision:
  • Refer true labels in a context-aware manner
A Representation Learning Approach

Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader Jesse James (e₁) in Missouri (e₂).

Gofraid (e₁) died in 989, said to be killed in Dal Riata (e₂).

Hussein (e₁) was born in Amman (e₂) on 14 November 1935.

Truth Discovery View

Relation Extraction View

Infer 'true' label

Training Truth Discovery model

Training Relation Extraction model

Labeling Functions

Relation Mentions

Vector Space

Interact through context representation

L₁ return born_in for <e₁, e₂, s> if BornIn(e₁, e₂) in KB
L₂ return died_in for <e₁, e₂, s> if DiedIn(e₁, e₂) in KB
L₃ return born_in for <e₁, e₂, s> if match(' * born in * ', s)
L₄ return died_in for <e₁, e₂, s> if match(' * killed in * ', s)

representation of proficient subset

representation of relation type died_in

representation of relation type born_in
A Representation Learning Approach

Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader Jesse James (e₁) in Missouri (e₂).

Gofraid (e₁) died in 989, said to be killed in Dal Riata (e₂).

Hussein (e₁) was born in Amman (e₂) on 14 November 1935.

\[ \text{Labeling Functions} \]

- \( \lambda_1 \) return \( \text{born in} \) for \( <e_1, e_2, s> \) if \( \text{BornIn}(e_1, e_2) \) in KB.
- \( \lambda_2 \) return \( \text{died in} \) for \( <e_1, e_2, s> \) if \( \text{DiedIn}(e_1, e_2) \) in KB.
- \( \lambda_3 \) return \( \text{born in} \) for \( <e_1, e_2, s> \) if \( \text{match(' * born in * ', s)} \).
- \( \lambda_4 \) return \( \text{died in} \) for \( <e_1, e_2, s> \) if \( \text{match(' * killed in * ', s)} \).

Heterogeneous Supervision generation

Relation Extraction View

- representation of relation mention
- representation of relation type \( \text{died in} \)
- representation of relation type \( \text{born in} \)

Relation Extraction View

- proficient subset
- representation of proficient subset

Relation Mention Representation

Relation Discovery View

- \( \mathcal{D} \)
- \( \mathcal{C}_1 \): Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader Jesse James (e₁) in Missouri (e₂).
- \( \mathcal{C}_2 \): Gofraid (e₁) died in 989, said to be killed in Dal Riata (e₂).
- \( \mathcal{C}_3 \): Hussein (e₁) was born in Amman (e₂) on 14 November 1935.

True Label Discovery

- \( \Lambda \)
- training Truth Discovery model
- training Relation Extraction model

Vector Space

Relation Mentions

Labeling Functions

Infer 'true' label

Context representation

Interact through context representation

Relation Extraction

representation of relation type \( \text{born in} \)

representation of relation type \( \text{died in} \)

representation of proficient subset

representation of relation mention

proficient subset
A Representation Learning Approach

Relation Mention Representation

Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader Jesse James ($e_1$) in Missouri ($e_2$).

Gofraid ($e_1$) died in 989, said to be killed in Dal Riata ($e_2$).

Hussein ($e_1$) was born in Amman ($e_2$) on 14 November 1935.

Truth Discovery View

Relation Extraction View

Interact through context representation

infer 'true' label

training Truth Discovery model

training Relation Extraction model

Labeling Functions

$\lambda_1$ return born_in for $<e_1,e_2,s>$ if BornIn($e_1,e_2$) in KB

$\lambda_2$ return died_in for $<e_1,e_2,s>$ if DiedIn($e_1,e_2$) in KB

$\lambda_3$ return born_in for $<e_1,e_2,s>$ if match( ' * born in * ', s)

$\lambda_4$ return died_in for $<e_1,e_2,s>$ if match( ' * killed in * ', s)

Heterogeneous Supervision generation

representation of proficient subset

representation of relation mention

proficient subset

representation of relation type died_in

representation of relation type born_in

Relation Extraction

True Label Discovery
Relation Mention Representation

- Text Feature Extraction
- Text Feature Representation
- Relation Mention Representation

**Mapping from Text Embedding to Relation Mention Embedding:**

\[
\text{tanh}(W \cdot \frac{1}{|f_i|} \sum_{f_i \in f_c} v_i)
\]

**Example Text:**

Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader Jesse James (\(e_1\)) in Missouri (\(e_2\)).

Gofraid (\(e_1\)) died in 989, said to be killed in Dal Riata (\(e_2\)).

Hussein (\(e_1\)) was born in Amman (\(e_2\)) on 14 November 1935.
Text Feature Extraction

We adopted texture features, POS-tagging and **brown clustering** to extract features.

C3: **Hussein** was born in **Amman** on 14 November 1935

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entity mention (EM) head</td>
<td>Syntactic head token of each entity mention</td>
<td>“HEAD_EM1_Hussein”, ...</td>
</tr>
<tr>
<td>Entity Mention Token</td>
<td>Tokens in each entity mention</td>
<td>“TKN_EM1_Hussein”, ...</td>
</tr>
<tr>
<td>Tokens between two EMs</td>
<td>Tokens between two EMs</td>
<td>“was”, “born”, “in”</td>
</tr>
<tr>
<td>Part-of-speech (POS) tag</td>
<td>POS tags of tokens between two EMs</td>
<td>“VBD”, “VBN”, “IN”</td>
</tr>
<tr>
<td>Collocations</td>
<td>Bigrams in left/right 3-word window of each EM</td>
<td>“Hussein was”, “in Amman”</td>
</tr>
<tr>
<td>Entity mention order</td>
<td>Whether EM 1 is before EM 2</td>
<td>“EM1_BEFORE_EM2”</td>
</tr>
<tr>
<td>Entity mention distance</td>
<td>Number of tokens between the two EMs</td>
<td>“EM_DISTANCE_3”</td>
</tr>
<tr>
<td>Body entity mentions numbers</td>
<td>Number of EMs between the two EMs</td>
<td>“EM_NUMBER_0”</td>
</tr>
<tr>
<td>Entity mention context</td>
<td>Unigrams before and after each EM</td>
<td>“EM_AFTER_was”, ...</td>
</tr>
<tr>
<td>Brown cluster (learned on $D$)</td>
<td>Brown cluster ID for each token</td>
<td>“BROWN_010011001”, ...</td>
</tr>
</tbody>
</table>

Diagram showing the mapping from text embedding to relation mention embedding.
Text Feature Representation

• Leverage features’ co-occurrence information to learn the representation, and help the model generalize better.

• Loss function of this part:

\[
J_E = \sum_{c \in C} \left( \log \sigma(v_i^T v_j^*) \right) - \sum_{k=1}^{V} E_{f_{k} \sim P} \left[ \log \sigma(-v_i^T v_{k}^*) \right]
\]

co-occurrence here refers to features occur in the same relation mention instead of the same shifting window

Feature embedding for feature \( f_i \)

Negative sampling
Relation Mention Representation

• Here, we adopted the bag-of-features assumption, and add transformation weights to allow representation of relation mention and features to be in different semantic space.

\[ z_c = g(f_c) = \tanh\left( W \cdot \frac{1}{|f_c|} \sum_{f_i \in f_c} v_i \right) \]
Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader Jesse James (e1) in Missouri (e2).

Goerlaid (e1) died in 989, said to be killed in Dal Riata (e2).

Hussein (e1) was born in Amman (e2) on 14 November 1935.
True label discovery

• Assume:
  • A labeling function would annotate similar instances with the same reliability
True label discovery

• Assume:
  • A labeling function would annotate similar instances with the same reliability

Context Information: $z$

for each labeling function, there exists a proficient subset, containing instances that it can precisely annotate.
True label discovery

• How to decide which label is correct?

  • Probability model and maximum likelihood estimate

  Corresponding to our assumption and setting

  Identify the true label
True label discovery

• Probability Model:
  • Describing the generation of Heterogeneous Supervision?
  • Different from crowdsourcing. E.g., ONE worker may annotate:
    • ("Obama", "USA", Obama was born in Honolulu, Hawaii, USA as he has always said)
      • Born-in
      • President-of
      • Citizen-of
      • ...
  Exists some randomness

Cautious Worker

Careless Worker
True label discovery

• Probability Model:
  • Describing the generation of Heterogeneous Supervision?
  • Different from crowdsourcing. E.g., **ONE** worker may annotate:
    • ("Obama", "USA", Obama was born in Honolulu, Hawaii, USA as he has always said)
      • **Born-in**
      • **President-of**
      • **Citizen-of**
      • ...

• But **One** labeling function can only annotate **One** relation type:
  • Randomness exists in the correctness, not in the choice of relation type
True label discovery

• Describing the correctness of Heterogeneous Supervision

\[
\rho_{c,i} = \delta(o_{c,i} = o_c^*)
\]

observed annotation

underlying true label

correctness of annotation \(o_{c,i}\)

whether \(c\) belongs to the proficient subset of \(l_i\)
True label discovery

• Describing the correctness of Heterogeneous Supervision

\[ p(\rho_{c,i} = 1) = p(\rho_{c,i} = 1|s_{c,i} = 1) \times p(s_{c,i} = 1) + p(\rho_{c,i} = 1|s_{c,i} = 0) \times p(s_{c,i} = 0) \]

\[ p(s_{c,i} = 1) = \sigma(l_i^T \ast z_c) \]

\[ J_T = \sum_{o_{c,i} \in O} \log(\sigma(z_c^T l_i)\phi_1^{(\sigma_{c,i} = o_{c}^*)}(1 - \phi_1)^{\delta(\sigma_{c,i} \neq o_{c}^*)} + (1 - \sigma(z_c^T l_i))\phi_0^{(\sigma_{c,i} = o_{c}^*)}(1 - \phi_0)^{\delta(\sigma_{c,i} \neq o_{c}^*)}) \]
A Representation Learning Approach

Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader Jesse James (e₁) in Missouri (e₂).

Gofraid (e₁) died in 989, said to be killed in Dal Riata (e₂).

Hussein (e₁) was born in Amman (e₂) on 14 November 1935.

Relation Mention Representation

Labeling Functions

λ₁ return born_in for <e₁, e₂, s> if BornIn(e₁, e₂) in KB

λ₂ return died_in for <e₁, e₂, s> if DiedIn(e₁, e₂) in KB

λ₃ return born_in for <e₁, e₂, s> if match(" * born in * ", s)

λ₄ return died_in for <e₁, e₂, s> if match(" * killed in * ", s)

Relation Extraction View

Interact through context representation

infer 'true' label

Heterogeneous Supervision generation

Relation Extraction

true label discovery

training Truth Discovery model

training Relation Extraction model

representation of proficient subset

representation of relation mention

representation of relation type died_in

representation of relation type born_in

D

Relation Mentions

Vector Space
Relation Extraction

- Adopts soft-max as the relation extractor:

\[
p(r_i | z_c) = \frac{\exp(z_c^T t_i)}{\sum_{r_j \in R \cup \{\text{None}\}} \exp(z_c^T t_j)}
\]

- Loss function: KL-Divergence:

\[
J_R = - \sum_{c \in C_l} KL(p(\cdot | z_c) || p(\cdot | o_c^*))
\]
A Representation Learning Approach

Relation Mention Representation

D

Robert Newton "Bob" Ford was an American outlaw best known for killing his gang leader Jesse James (e1) in Missouri (e2).

Hussein (e2) was born in Amman (e2) on 14 November 1935.

Gofraid (e1) died in 989, said to be killed in Dal Riata (e2).

Relation Mentions

Labeling Functions

λ1 return born_in for <e1,e2, s> if BornIn(e1,e2) in KB

λ2 return died_in for <e1,e2, s> if DiedIn(e1,e2) in KB

λ3 return born_in for <e1,e2, s> if match(" * born in * ", s)

λ4 return died_in for <e1,e2, s> if match(" * killed in * ", s)

Vector Space

Heterogeneous Supervision generation

Relation Extraction View

Truth Discovery View

training Truth Discovery model

λ1 λ2 λ3

λ4

true

training Relation Extraction model

representation of relation type died_in

representation of relation type born_in

representation of relation mention

representation of proficient subset

Relation Extraction

Interact through context representation

infer 'true' label

D1 D2 D3

intermediate relation representation of relation mention

intermediate relation representation of relation type

intermediate relation type `true'

Heterogeneous Supervision generation

True Label Discovery
Model Learning

• Joint optimize three components

\[
\min_{W,v,v^*,l,t,o^*} J = -J_R - \lambda_1 J_E - \lambda_2 J_T
\]
\[
\text{s.t. } \forall c \in C_l, o^*_c = \arg\max_{o^*_c} J_T, z_c = g(f_c)
\]
ReHession

• Heterogeneous Supervision

• Our Solution: A Representation Learning Approach
  • Relation Mention Representation
  • True Label Discovery component
  • Relation Extraction component

• Experiments
Experiments

• 1. Relation extraction (with None) and Relation classification (without None):
  • NL: train relation extractor with all annotations
  • TD: train relation extractor with ‘true’ label inferred by Investment (compared true label discovery model)
### Experiments

<table>
<thead>
<tr>
<th>Method</th>
<th>Relation Extraction</th>
<th>Relation Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NYT Prec</td>
<td>NYT Rec</td>
</tr>
<tr>
<td>NL+FIGER</td>
<td>0.2364</td>
<td>0.2914</td>
</tr>
<tr>
<td>NL+BFK</td>
<td>0.1520</td>
<td>0.0508</td>
</tr>
<tr>
<td>NL+DSL</td>
<td>0.4150</td>
<td>0.5414</td>
</tr>
<tr>
<td>NL+MultiR</td>
<td>0.5196</td>
<td>0.2755</td>
</tr>
<tr>
<td>NL+FCM</td>
<td>0.4170</td>
<td>0.2890</td>
</tr>
<tr>
<td>NL+CoType-RM</td>
<td>0.3967</td>
<td>0.4049</td>
</tr>
<tr>
<td>TD+FIGER</td>
<td>0.3664</td>
<td>0.3350</td>
</tr>
<tr>
<td>TD+BFK</td>
<td>0.1011</td>
<td>0.0504</td>
</tr>
<tr>
<td>TD+DSL</td>
<td>0.3704</td>
<td>0.5025</td>
</tr>
<tr>
<td>TD+MultiR</td>
<td><strong>0.5232</strong></td>
<td>0.2736</td>
</tr>
<tr>
<td>TD+FCM</td>
<td>0.3394</td>
<td>0.3325</td>
</tr>
<tr>
<td>TD+CoType-RM</td>
<td>0.4516</td>
<td>0.3499</td>
</tr>
<tr>
<td>REHession</td>
<td>0.4122</td>
<td><strong>0.5726</strong></td>
</tr>
</tbody>
</table>

Table 6: Performance comparison of relation extraction and relation classification
Experiments

• 2. Effectiveness of proposed true label discovery component:
  • Ori: with proposed context-aware true label discovery component
  • LD: with Investment (compared true label discovery model)

<table>
<thead>
<tr>
<th>Dataset &amp; Method</th>
<th>Prec</th>
<th>Rec</th>
<th>F1</th>
<th>Acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wiki-KBP</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ori</td>
<td>0.3677</td>
<td>0.4933</td>
<td>0.4208</td>
<td>0.7277</td>
</tr>
<tr>
<td>TD</td>
<td>0.3032</td>
<td><strong>0.5279</strong></td>
<td>0.3850</td>
<td>0.7271</td>
</tr>
<tr>
<td>NYT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ori</td>
<td><strong>0.4122</strong></td>
<td>0.5726</td>
<td><strong>0.4792</strong></td>
<td><strong>0.8381</strong></td>
</tr>
<tr>
<td>TD</td>
<td>0.3758</td>
<td>0.4887</td>
<td>0.4239</td>
<td>0.7387</td>
</tr>
</tbody>
</table>

Table 7: Comparison between REHESSION (Ori) and REHESSION-TD (TD) on relation extraction and relation classification
## Case Study

<table>
<thead>
<tr>
<th>Relation Mention</th>
<th>REHession</th>
<th>Investment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ann Demeulemeester (born 1959, Waregem, Belgium) is a ...</td>
<td>born-in</td>
<td>None</td>
</tr>
<tr>
<td>Raila Odinga was born at ..., in Maseno, Kisumu District, ...</td>
<td>born-in</td>
<td>None</td>
</tr>
<tr>
<td>Ann Demeulemeester (elected 1959, Waregem, Belgium) is a ...</td>
<td>None</td>
<td>None</td>
</tr>
<tr>
<td>Raila Odinga was examined at ..., in Maseno, Kisumu District, ...</td>
<td>None</td>
<td>None</td>
</tr>
</tbody>
</table>

Table 8: Example output of true label discovery. The first two relation mentions come from Wiki-KBP, and their annotations are \{born-in, None\}. The last two are created by replacing key words of the first two. Key words are marked as bold and entity mentions are marked as Italics.
Thank You

Q & A