Type-augmented Relation Prediction in Knowledge Graphs

Motivation

- Leverage prior type information to improve relation prediction performance
- Relation Prediction in Knowledge Graphs:
  - o (Helen Mirren, ?, Chiswick)
- Prior Knowledge: type information of entities/relations
  - o Helen Mirren is a person/award_winner/actor/
  - o Person, place_of_birth, location

Type Information Encoding

- We encode the type information as prior probabilities by considering hierarchical structures among types
- Type sets usually have an underlying hierarchy, such as the structure among types (actor, award_winner, person):
  - $H_1 = \{\text{person}, \text{actor}\}$
  - $H_2 = \{\text{person}, \text{award_winner}\}$
  - $H_3 = \{\text{person}\}$
- Hierarchy-based type weights
  - We define hierarchy-based type weights to assign different weights to types based on their locations in the hierarchy
  - We hypothesize that types of more specific semantic meaning are more helpful, and higher weights are automatically assigned to these types
  - For example, given three hierarchies $H_1$, $H_2$, and $H_3$, we have type weights:
    - $w_1(\text{person}) = \min(0.27, 0.27, 1) = 0.27$
    - $w_2(\text{actor}) = 0.73$
    - $w_3(\text{award_winner}) = 0.73$
- Type-based prior probability
  - Given a triple $(e_u, r, e_v) \in \mathcal{G}$, we define two similarity score $s(e_u, r)$ and $s(e_v, r)$ based on the correlation between type sets
  - The prior probability $p(r \mid (e_u, e_v, \mathcal{R}))$ is then defined as
    $$p(r \mid (e_u, e_v, \mathcal{R})) = \frac{s(e_u, r) s(e_v, r)}{\sum_{r' \in \mathcal{R}} s(e_u, r') s(e_v, r')}$$
- The higher the correlation between type sets, the higher the prior probability of the relation

Embedding-based Models

- Embedding-based models learn representations of relations and entities by minimizing the distance $f_r(e_u, e_v)$ in a continuous embedding space
- Given the learned embeddings, we compute the likelihood by taking the exponential
  $$p(e_u, e_v \mid r) = \exp(f_r(e_u, e_v))$$
  - The lower the distance, the lower the likelihood

Conclusions

- We achieve significantly better performance by leveraging type information compared to SoTAs on four benchmark datasets
- Our proposed approach is effective in integrating type information
- In the paper, we also show that our method is more data efficient. Through cross-dataset evaluation, we show that type information extracted from a specific dataset can generalize well to different datasets