Challenging the Assumption of Structure-based embeddings in Few- and Zero-shot Knowledge Graph Completion

Filip Cornell\textsuperscript{1,3}, Chenda Zhang\textsuperscript{2}, Šarūnas Girdzijauskas\textsuperscript{1}, Jussi Karlberg\textsuperscript{2}

\textsuperscript{1}KTH Royal Institute of Technology, \textsuperscript{2}Carnegie Mellon University, \textsuperscript{3}Gavagai

Introduction

- Previous work in relational Few- and Zero-shot KG completion trains structure-based embeddings and then uses a complex neural architecture to infer missing links.
- What if we instead use textual embedding representations of the entities and relations, instead of structural?
- Textual embeddings are not limited to the knowledge in the graph, but to the knowledge it has been trained on!

Issue: Textual descriptions not available for some benchmark datasets!

Solution: We collect textual descriptions from existing sources to compile additional data for benchmark datasets.

Result: Improvements in inference in Zero- and Few-shot Knowledge Graph completion. Better improvements the Wiki datasets, as descriptions are more informative.

Textual data collected is available at https://bit.ly/3MHYCxV

Few- and ZS KG completion

Data collection

4 benchmark datasets

Wiki datasets
- Gather label, instance & description from the Wikidata API
- Translate into English

NELL datasets
- Entity descriptions were not found
- Entities defined
- Relation descriptions available from the NELL project\textsuperscript{1}

Note: Zero-shot datasets are subsets of the Few-shot datasets - facilitates collection!

Creating embeddings

SIF\textsuperscript{3}
- Weighs the word embeddings of the words in the descriptions smartly, followed by removing the first five principal components to reduce noise.
- 100-dimensional embeddings
- Pretrained embeddings used: glove-wiki-gigaword-100

Sentence-BERT\textsuperscript{4}
- Embeds descriptions as contextual sentences
- Transformer-based architecture, trained to represent semantic meanings of sentences
- 384 dimensions
- Pretrained model used: paraphrase-MiniLM-L12-v2

Experiments

Compare against structure-based pre-trained embeddings
- Run models under same settings with different representations of the entities — compare
- Compare against TransE and DistMult

Few-shot:
- We use the PAAN\textsuperscript{5} architecture
- Stability issues using Sentence-BERT - only SIF used here.

Zero-shot: We use the OntoZSL\textsuperscript{6} and ZS-GAN\textsuperscript{7}, two previous architectures.

Note: No architectural modifications! We only replace inputs!

Results

- Improvement in both Few- and Zero-Shot scenario
- Improvements are smaller for NELL — descriptions are shorter and less informative

References


Acknowledgments

This work was partially supported by the Wallenberg AI Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. Many thanks also to Martin Isaksson, Fredrik Olsson, Tobias Norlund, and Sofie Verrewaere for valuable input and advice.

Creating embeddings from text

Replacing structural embeddings as input

Our process

Available before

What we collect

Note

Few-shot

Zero-shot KG completion

Relations to predict for have not been seen in the graph before!

Data collection

Creating embeddings

Replacing structural embeddings as input

Dataset | Data type | ID | Label | Description
--- | --- | --- | --- | ---
QB3115 | Sentence: video game | G0840 | Payrin dancer | Payrin dancer, the town of Payrin, France
Wiki | Entities | Q63115 | Fantasy World Directory | Fantasy World Directory
Wiki | Entities | Q17580427 | Church of St Mary | Church of St Mary
Wiki | Relations | P37 | Official language | Official language
Wiki | Relations | P112 | Physically interacts with | Physically interacts with
Wiki | Relations | P2382 | Place of attention | Place of attention
NELL | Entities | concept:company:keva | Knowability | Knowability
NELL | Relations | concept:includable:matchability | Concept-Includable | Concept-Includable

FAQ

What do we collect

- Few- and Zero-shot KG completion
- Relations to predict

What we collect

- Textual data collected is available at https://bit.ly/3MHYCxV
- Wikipedia
- NELL

What we collect

- Embeds descriptions as contextual sentences
- Transformer-based architecture, trained to represent semantic meanings of sentences
- 384 dimensions
- Pretrained model used: paraphrase-MiniLM-L12-v2

Textual data collected is available at https://bit.ly/3MHYCxV

Sentence-BERT

- Weighs the word embeddings of the words in the descriptions smartly, followed by removing the first five principal components to reduce noise.
- 100-dimensional embeddings
- Pretrained embeddings used: glove-wiki-gigaword-100