Experimenting Embeddings with Graph Neural Networks for Knowledge Graphs using RDF Reification

Master 2 research internship. GDD Team. LS2N Laboratory.
Keywords: graph neural networks, embeddings, knowledge graphs, RDF reification.
Required skills: basis on Machine Learning, Knowledge Graphs (RDF), Linux commands, Python, and Docker.
Contacts: Patricia.Serrano-Alvarado@univ-nantes.fr, Manoe.Kieffer@univ-nantes.fr, Khalid.Belhajjame@lamsade.dauphine.fr

Introduction and context

The context of this work is the CLARA project [CLARABench][1]. In this internship, we are interested in analyzing knowledge graphs using deep learning methods. Since their introduction, deep learning models have been at the center of attention. The latest examples are the Large Language Models (LLMs) and in particular the transformer model [Transformer] used by ChatGPT. These models are built for tasks such as chatbots, conversational AIs, or sentiment analysis of texts. However, deep learning models have proven to be very efficient for other tasks, like the Convolutional Neural Networks (CNNs) for image recognition. This efficiency has also been proven for analyzing knowledge graphs with Graph Neural Networks (GNN) [SurveyGNN]. GNN models are particularly efficient for tasks such as link prediction, entity classification, or k-nearest neighbours.

Knowledge graphs represent statements as triples (head, relation, tail). Each triple is a fact stating a relation between two entities. Statements about statements, also called statement-level annotations, are increasingly used. They allow specifying that a statement is true under a particular context. Context can concern temporal aspects, provenance, trust values, scores, weights, etc. RDF reification allows expressing statement-based annotations in a generic way. We believe that current GNN models are not suited for processing annotations. GNN models do not support RDF reification because it may introduce noise that would reduce the quality of the results.

The goal of this internship is to show the limits of existing GNN models in the presence of RDF reification and to propose a new model that efficiently integrates RDF annotations.

Objectives

You will participate in research work alongside a PhD student in the following tasks:
- **Defining and running an experimental protocol.** You will conceptualize and run an experimental protocol to put into light the impact of reification on GNN models. This will require the understanding of (a) several models (e.g., R-GCN [RGCN], HypE [HypE], RDF2vec [RDF2vec]), (b) the different reification approaches and their impact on GNN models (standard reification [Standard], n-ary relations [N-ary],

---

[1] https://project.inria.fr/clara/

RDF-star [RDF-star]), and (c) how the impact of RDF reification on the GNN models can be measured. The obtained results will be the baseline for the next task.

- **Creating a new GNN model.** You will help in contributing a GNN model that better integrates RDF reification. The model will be inspired by the message-passing algorithm used in GNNs such as R-GCN and it should be able to adapt to RDF reification. The result of this approach should be compared to the baseline previously obtained.

Bibliography

[CLARABench] Manoé Kieffer, Ginwa Fakih, and Patricia Serrano-Alvarado: Evaluating Reification with Multi-valued Properties in a Knowledge Graph of Licensed Educational Resources. SEMANTiCS, 2023.

[Transformer] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, Illia Polosukhin: Attention is all you need. Advances in neural information processing systems, 2017.

[SurveyGNN] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, Philip S. Yu: A comprehensive survey on graph neural networks. IEEE transactions on neural networks and learning systems, 2020.

[Standard] Manola F, Miller E, McBride B, et al. RDF primer. W3C recommendation. 2004. Available from: https://www.w3.org/TR/rdf-primer/.

[N-ary] Pat Hayes, Chris Welty: Defining N-ary Relations on the Semantic Web. 2006. Available from: https://www.w3.org/TR/swbp-n-aryRelations/.

[RDF-star] Olaf Hartig: Foundations of RDF* and SPARQL* (An alternative approach to statement-level metadata in RDF). International Workshop on Foundations of Data Management and the Web (AMW), 2017.

[RGCN] Michael Schlichtkrull, Thomas N. Kipf, Peter Bloem, Rianne van den Berg, Ivan Titov, Max Welling: Modeling relational data with graph convolutional networks. 15th International Conference on Extended Semantic Web Conference, 2018.

[HypE] Bahare Fatemi, Perouz Taslakian, David Vazquez, David Poole: Knowledge hypergraphs: Prediction beyond binary relations. Proceedings of the Twenty-Ninth International Conference on International Joint Conferences on Artificial Intelligence. 2021.

[RDF2vec] Petar Ristoski, Heiko Paulheim: RDF2vec: RDF graph embeddings for data mining. 15th International Semantic Web Conference, 2016.