

Knowledge Graph Inference using Tensor Embedding

Ankur Padia¹*, Konstantinos Kalpakis², Francis Ferraro², Tim Finin²

¹Philips Research North Americas, Cambridge, MA 02141, USA

²University of Maryland, Baltimore County, Baltimore, MD 21250, USA

ankur.padia@philips.com, {kalpakis, ferraro, finin}@umbc.edu

1 Introduction

This extended abstract describes work reported in the Journal of Web Semantics (Padia et al. 2019).

Axiom based inference provides a clear and consistent way of reasoning to add more information to a knowledge graph. However, constructing a set of axioms is expensive and requires domain expertise, time, and money. It is also difficult to reuse or adapt a set of axioms to a knowledge graph in a new domain or even in the same domain but using a slightly different representation approach.

Representation learning (Bengio, Courville, and Vincent 2013) introduces a way to augment or even replace manually constructed ontology axioms and rules by using knowledge graph instances to discover common patterns and then apply them to suggest changes to the graph. The changes are often in the form of adding missing types and relations, but can also include schema modifications, removing incoherent instances, merging sets of instances describing the same real-world entity, or adding relation probabilities.

One popular approach for representation learning is based on learning how to embed a graph’s entities and relations into a real-valued vector space, allowing both to be represented by dense, real-valued vectors. The entity and relation embeddings can be learned either independently or jointly, and then used to predict additional relations that are missing. Jointly learning the embeddings allows each to enhance the other (Nickel, Tresp, and Krieger 2011).

There are several models that learn embedding of entities and relations to perform inference. Some use existing schemas (Krompass, Baier, and Tresp 2015) to regularize the quality of the embedding. These often give better performance compared to those that do not use schemas, as in Nickel (2011). However, schema-based embedding methods suffer from the above-mentioned limitations. We have developed a family of novel methods that improves the quality of the embeddings without using pre-defined schemas (Padia et al. 2019; Padia 2019a).

We divide statistical models that infer additional knowledge graph facts into two categories: *link ranking* systems that provide a ranked list of the top-K candidates and *link*

prediction or classification systems that determine if a new fact holds or not. There are several approaches to link ranking (Socher et al. 2013), all of which involve an auxiliary problem of determining a threshold, either globally or per relation, that separates plausible from implausible relations. Since we are only interested in extending a knowledge graph with relations that are likely to hold (what we call facts), we designed an approach to solve it directly. Thus we have the fact or link prediction task: given a knowledge graph, learn a model that can classify relation instances that are very likely to hold. This task is more specific than link ranking and more directly solves an important problem.

We improve the quality of the relation and entities representations using data-driven constraints, hence our approach can be used when ontological axioms are not available or are expensive to create. We measured the quality of the learned embeddings by comparing our approaches with previous non-schema based methods as well as with neural models and found improvement ranging from 5% to 50%. We demonstrated its broad applicability using eight real-world data sets covering human language, medical data, and general world knowledge that are available online (Padia 2019b). This work makes three main contributions: it (1) provides a family of representation learning algorithms and an extensive analysis on eight datasets; (2) yields better results than existing tensor and neural models; and (3) includes a provably convergent factorization algorithm.

We use the initial knowledge graph to pre-compute a similarity matrix C for the relations that will help constrain the learning of embeddings. To better understand the idea, consider the WordNet knowledge graph where entities are words that are connected with relations like *hypernym* and *similar*. The graph has no schema and there are many missing relations. We create a similarity matrix quantifying the similarity between two relations as the number of overlapping words. The cells in Figure 1 show the number of subjects or objects that are shared by the relations, with darker cells indicating a smaller overlap.

2 Models Using Axiomatic-like constraints

We represent a multi-relational knowledge graph of N_r binary relations and N_e entities by the order-3 tensor \mathcal{X} of dimension $N_e \times N_e \times N_r$. This binary tensor is often both very large and sparse. Our goal is to construct dense, infor-

*This work was done when the first author was doing Ph.D. at the University of Maryland, Baltimore County and before joining Philips Research North America.

