John (Ziyuan) Zhou

Week 5 Summary

UW ID: 2150115

In the article “Beta embeddings for Multi-Hop Logical Reasoning in Knowledge Graphs”, the authors point out the problem in Artificial intelligence, which is to perform complex multi-hop logical reasoning over the facts captured by the knowledge graph (KG). There are two main challenges to this problem of solving the KG reasoning – the scale and completeness of KG. The recent approach mainly focuses on designing neural logical operators and embed queries iteratively by executing logical operations according to the query computation graph. The advantage is that it no longer needs to track all the intermediate entities, that is, can use the nearest neighbor search in the embedding space to discover answer quickly. However, the drawback is that it only supports existential positive first order (EPFO) queries, which is a subset of FOL queries without negation. Furthermore, it also cannot naturally model uncertainly. To solve these problems, the authors propose Beta Embedding (BETAE), which is a probabilistic embedding framework for answering arbitrary FOL queries over KG. The main idea is to model both the entities and queries by probabilistic distribution with bounded support, and this design has the following advantages:

* Capturing the uncertainty of the queries efficiently.
* Supporting full first order logic: existential quantification, conjunction, disjunction, and negation.
* Conjunction and negation naturally correspond to the real operations and captures several properties of first order logic.
* Can handle complete set of FOL operators and thus supporting arbitrary FOL queries.

Besides this, in order to answer a query using the computation graph, the authors design the probabilistic logical operators for the BETAE, which are probabilistic projection, probabilistic intersection, and probabilistic negation. For the experiment and result, the authors test the performance of BETAE on standard KG dataset. Compared to the recent approach, BETAE achieves the state-of-the-art performance in handling arbitrary conjunctive queries with a relative increase of accuracy by up to 25.4%.