Summer Project 2020

Reinforcement Knowledge Graph Reasoning for Explainable Recommendation

Aditi Goyal*
Computer Science and Engineering, IIT Kanpur

Rahul Sethi†
Mathematics and Scientific Computing, IIT Kanpur

Somya Lohani‡
Computer Science and Engineering, IIT Kanpur

Vansh Bansal§
Computer Science and Engineering, IIT Kanpur

Mentored by:
Ishika Gupta
Computer Science and Engineering & Chemical Engineering, IIT Kanpur

Contents

*190057, aditi@iitk.ac.in
†190668, sethir@iitk.ac.in
‡190848, somyaloh@iitk.ac.in
§190941, vanshb@iitk.ac.in

Abstract

Knowledge Graphs connect various types of information related to items into a unified space. Different paths connecting entity pairs often carry relations of different semantics, and PGPR* (Policy Guided Path Reasoning) models these with the help of high quality user and item representations generated using the TransE[1] graph embedding scheme.

I. Nomenclature

- G - discounted cumulative reward from s to s_T
- s_T - terminal state
- γ - discount factor
- $\hat{v}(s)$ - value network, used as baseline for REINFORCE
- $\tilde{A_u}$ - user conditional pruned action space

II. Introduction

A. Overview

This project:

- highlights the importance of KGs to define and interpret the process of recommendation.
- proposes an RL-based approach (with soft rewards, a multi-hop scoring function and action pruning)
- imposes a beam search algorithm to sample diverse reasoning paths and items for recommendation.
- evaluates this method on four Amazon datasets to get explicit reasoning behind the predicted paths.

B. Problem Formulation

Goal: Given a user u, find a set of candidate items i_n and the corresponding reasoning paths $p_n(u, i_n)$

- Entity Set E
- Relation Set R
- Users U, Items I such that $U \cap I = \phi$ and $U, I \subseteq E$

III. Preliminaries

A. Datasets

We've evaluated the PGPR Model[2] on the following four datasets:

1) Amazon Beauty
2) Amazon Cell Phones
3) Amazon CDs & Vinyl
4) Amazon Clothing

B. Metrics

Model evaluation is done in terms of four representative top-N recommendation measures. These ranking metrics are computed based on the top-10 predictions for every user in the test set.

I. **NDCG**

Normalized Discounted Cumulative Gain is a popular method for measuring the quality of a set of search results. It asserts the following:

- Cumulative Gain - The usefulness of very relevant results, somewhat relevant results and irrelevant results decreases in that order
- Discounting - Relevant results are more useful when they appear earlier in the set of results
- Normalization - The result of the ranking should be irrelevant to the query performed

*Code available here: https://github.com/knighterudite/PGPR

II. **Hit Rate**
Metric to evaluate top-N recommendations to a particular user, based on leave-one-out cross validation method.

III. **Recall** (Sensitivity)
Fraction of relevant instances retrieved among total relevant instances

IV. **Precision** (Positive Predictive Value)
Fraction of relevant instances among retrieved instances

C. Optimization

Goal: To learn a stochastic policy π that maximizes the expected cumulative reward for a particular initial user u

$$\mathrm{J}(\theta) = \mathbb{E}_{\pi}\left[\sum_{t=0}^{T-1} \gamma^{t} \mathrm{R}_{t+1} | u\right]$$

$$\nabla_{\theta}\mathrm{J}(\theta) = \mathbb{E}_{\pi}\left[\nabla_{\theta} \log \pi_{\theta}(\cdot|\mathrm{s}, \mathrm{A_u})(\mathrm{G} - \hat{v}(s))\right]^{\dagger}$$

IV. Experiments

A. Framework

The implementation‡ consists of four stages§:

1) Data pre-processing step
2) Training TransE graph embeddings for entities and relations
3) Training RL agent
4) Testing - Model Evaluation

B. Results

Dataset: Beauty			
NDCG	Recall	Hit Rate	Precision
5.449	8.324	14.401	1.707

Dataset: Cell Phones			
NDCG	Recall	Hit Rate	Precision
5.042	8.416	11.904	1.274

Dataset: CDs & Vinyl			
NDCG	Recall	Hit Rate	Precision
5.590	7.569	16.886	2.157

Dataset: Clothing			
NDCG	Recall	Hit Rate	Precision
2.858	4.834	7.020	0.728

V. Conclusion

The model not only achieves outstanding recommendation results, but also supports the same with an interpretable causal inference procedure. The PGPR approach is a flexible graph reasoning framework and can be extended to many other graph-based tasks such as product search and social recommendation.

† Detailed discussion here: https://github.com/knighterudite/PGPR/blob/master/docs/Actor_Critic.pdf
‡ Demo run here: https://github.com/knighterudite/PGPR/blob/master/demo_run_beauty.ipynb
§ Details here: https://github.com/knighterudite/PGPR/blob/master/docs/code_detail.pdf

VI. Appendix

1. NDCG

CG = $\sum_{k=1}^{n} r_k$ (where r = relevance value for a particular result and CG = Cumulative Gain)
DCG rectifies CG by discounting the results that appear later, i.e. DCG = $\sum_{k=1}^{n} \frac{r_k}{\log(k+1)}$
$\text{NDCG}_k = \frac{\text{DCG}_k}{\text{iDCG}}$, the DCGs are normalized across queries by dividing by the best/ideal DCG.

References

[1] Antoine Bordes, Nicolas Usunier, Alberto Garcia-Duran, Jason Weston, Oksana Yakhnenko, "Translating Embeddings for Modeling Multi-relational Data," *Advances in Neural Information Processing Systems*, 2013, pp. 2787–2795.

[2] Yikun Xian, Zuohui Fu, S. Muthukrishnan, Gerard de Melo, Yongfeng Zhang, "Reinforcement Knowledge Graph Reasoning for Explainable Recommendation," *Conference on Research and Development in Information Retrieval (SIGIR '19)*, ACM, New York, NY, USA, 2019.