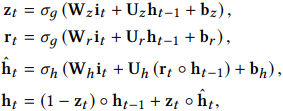
John (Ziyuan) Zhou

Week 7 Summary

UW ID: 2150115

In the article “Interactive Recommender System via Knowledge Graph-enhanced Reinforcement Learning”, the authors make the first attempt to leverage KG for reinforcement learning in Interactive Recommender System (IRS). Recently, reinforcement learning (RL) was widely used in IRS, and RL methods share a common issue of sample efficiency – huge amount of interaction data is required for training which caused by the sparse user responses and the large action space consisting of many candidate items. Furthermore, it is not feasible to collect much data with explorative policies in online environment, which might will harm user experience. Thus, they propose Knowledge Graph enhanced Q-learning framework for interactive Recommendation (KGQR), which is a novel architecture that extends deep Q-network (DQN). Specifically, they integrate graph learning and sequential decision making to facilitate knowledge in KG and pattern mining in IRS. By leveraging prior knowledge in KG in both candidate selection and the learning of user preference form sparse user feedback, KGQR can improve sample efficiency of RL-based IRS models. KGQR model contains four main components – graph convolution module, state representation module, candidate selection module, and Q-learning network module.

* Graph convolution module and state representation module: The graph convolutional network (GCN) is used to recursively propagate embeddings along the connectivity of items and learn the embeddings along the connectivity of items and learn the embeddings of all entities on the graph. The computation of node’s representation in a single graph convolutional embedding layer has two steps: aggregation and integration. In each layer, first, aggregating the representations of the neighboring nodes of a given node h: . Second, integrating the neighborhood representation with h’s representation as . After k-hop graph convolutional embedding layer, each clicked item is then converted into . Recurrent neural network (RNN) with gated recurrent unit (GRU) is used as the network cell to aggregate user’s historical behaviors and distill user’s state .
  + 
  + 
* Candidate selection module: A sampling based on the k-hop neighborhood in KG is used. In each timestep , the user’s historical interacted items serve as the seed set . The k-hop neighborhood set starting from the seed entities is denoted as . The candidate action set for the current user state is defined as .
* Q-learning network module: Q-network is used to combine information of candidate sets and improve the recommendation policy for the interactive recommendation process. Thus, the authors implement a DQN with dueling-Q and double-Q techniques to model the expected long-term user satisfaction form the current user state as well as to learn the optimal strategy. Dueling technique is used to reduce the approximation variance and stabilize training. The Q-value can be computed as . . During the model training, they sample mini-batch of experiences and minimize the mean-square loss function to improve the Q-network, defined as . is the target value based on the optimal Q, defined as . To alleviate the overestimation problem in original DQN, the target network Q along with the online network Q is utilized. So, the target value is changed to . .

For the experiment and result, the authors used two real-world datasets to evaluate the KGQR framework, which demonstrate the superiority of their approach with significant improvements against state-of-the-arts.