

先端科学技術研究科 修士論文要旨

所属研究室 (主指導教員)	知能コミュニケーション (中村 哲 (教授))		
学籍番号	2111437	提出日	令和 5年 7月 24日
学生氏名	WU DONGMING		
論文題目	Pretraining Language Models Using Hyperlinks and Entity Relations for Multi-hop Question Answering		
要旨			
<p>Multi-hop question answering is a challenging question answering task that requires comprehending and reasoning with information from more than one document to find an answer for a question. Several studies discovered that a higher performance on this task can be achieved by utilizing hyperlinks between documents or knowledge graphs during fine-tuning or language models pretraining. In this work, we investigate whether the performance on the multi-hop question answering can further be boosted by pretraining language models with (1) multiple types of hyperlinks or, with (2) entity relations in knowledge graphs. Specifically, for (1) we pretrain a language model with hyperlinked document pairs using two types of hyperlinks in Wikipedia, one type of which have not been used in the previous work. For (2), we extract entity relations from a publicly available knowledge graph, Wikidata, and train language models to recognize or predict the words which are corresponding to entity relations in Wikipedia articles. We pretrained BERT-based language models using the methods (1) and (2) separately and then fine-tuned each pretrained language model on HotpotQA, a multi-hop question answering dataset. Our experimental results showed that (1) the language model pretrained with hyperlinked document pairs using multiple types of hyperlinks achieved a higher performance on the downstream multi-hop question answering than those of existing language models. However, (2) the language models pretrained with entity relations in the knowledge graph could not improve the performance on downstream multi-hop question answering.</p>			