cases	doc_1	doc_2	decision	id
			DUPLICATES	334

doc_1	
authors	• Yuki Arase • Jun'ichi Tsujii
title	Transfer Fine-Tuning: A BERT Case Study
publication_date	2019-09-02 00:00:00
source	SupportedSources.OPENALEX
journal	arXiv (Cornell University)
volume	
doi	10.48550/arxiv.1909.00931
urls	• https://openalex.org/W4288122335 • https://doi.org/10.48550/arxiv.1909.00931 • http://arxiv.org/pdf/1909.00931
id	id-2058114509324601361
abstract	
versions	

doc_2	
authors	• Junichi Tsujii • Yuki Arase
title	Transfer Fine-Tuning: A BERT Case Study
publication_date	2019-09-03 00:00:00
source	SupportedSources.PAPERS_WITH_CODE
journal	
volume	
doi	
urls	• https://arxiv.org/pdf/1909.00931v1.pdf • https://github.com/yukiar/TransferFT • https://aclanthology.org/D19-1542.pdf
id	id-8731650874298029426
abstract	A semantic equivalence assessment is defined as a task that assesses semantic equivalence in a sentence pair by binary judgment (i.e., paraphrase identification) or grading (i.e., semantic textual similarity measurement). It constitutes a set of tasks crucial for research on natural language understanding. Recently, BERT realized a breakthrough in sentence representation learning (Devlin et al., 2019), which is broadly transferable to various NLP tasks. While BERT's performance improves by increasing its model size, the required computational power is an obstacle preventing practical applications from adopting the technology. Herein, we propose to inject phrasal paraphrase relations into BERT in order to generate suitable representations for semantic equivalence assessment instead of increasing the model size. Experiments on standard natural language understanding tasks confirm that our method effectively improves a smaller BERT model while maintaining the model size. The generated model exhibits superior performance compared to a larger BERT model on semantic equivalence assessment tasks. Furthermore, it achieves larger performance gains on tasks with limited training datasets for fine-tuning, which is a property desirable for transfer learning.
versions	