

cases	doc_1		doc_2		decision	id
	authors	<ul style="list-style-type: none">Wei WangBin BiMing YanChen WuZuyi BaoLiwei PengLuo Si	authors	<ul style="list-style-type: none">Wei WangBin BiMing YanChen WuZuyi BaoJiangnan XiaLiwei PengLuo Si	DUPLICATES	52
	title	StructBERT: Incorporating Language Structures into Pre-training for Deep Language Understanding	title	StructBERT: Incorporating Language Structures into Pre-training for Deep Language Understanding		
	publication_date	2019-08-01 00:00:00	publication_date	2019-09-27 00:00:00		
	source	SupportedSources.SEMANTIC_SCHOLAR	source	SupportedSources.INTERNET_ARCHIVE		
	journal	ArXiv	journal			
	volume	abs/1908.04577	volume			
	doi		doi			
	urls	<ul style="list-style-type: none">https://www.semanticscholar.org/paper/d56c1fc337fb07ec004dc846f80582c327af717c	urls	<ul style="list-style-type: none">https://web.archive.org/web/20200916135038/https://arxiv.org/pdf/1908.04577v3.pdf		
	id	id1477760616247244039	id	id-3009754849776420566		
	abstract	Recently, the pre-trained language model, BERT (and its robustly optimized version RoBERTa), has attracted a lot of attention in natural language understanding (NLU), and achieved state-of-the-art accuracy in various NLU tasks, such as sentiment classification, natural language inference, semantic textual similarity and question answering. Inspired by the linearization exploration work of Elman [8], we extend BERT to a new model, StructBERT, by incorporating language structures into pre-training. Specifically, we pre-train StructBERT with two auxiliary tasks to make the most of the sequential order of words and sentences, which leverage language structures at the word and sentence levels, respectively. As a result, the new model is adapted to different levels of language understanding required by downstream tasks. The StructBERT with structural pre-training gives surprisingly good empirical results on a variety of downstream tasks, including pushing the state-of-the-art on the GLUE benchmark to 89.0 (outperforming all published models), the F1 score on SQuAD v1.1 question answering to 93.0, the accuracy on SNLI to 91.7.	abstract	Recently, the pre-trained language model, BERT (and its robustly optimized version RoBERTa), has attracted a lot of attention in natural language understanding (NLU), and achieved state-of-the-art accuracy in various NLU tasks, such as sentiment classification, natural language inference, semantic textual similarity and question answering. Inspired by the linearization exploration work of Elman [8], we extend BERT to a new model, StructBERT, by incorporating language structures into pre-training. Specifically, we pre-train StructBERT with two auxiliary tasks to make the most of the sequential order of words and sentences, which leverage language structures at the word and sentence levels, respectively. As a result, the new model is adapted to different levels of language understanding required by downstream tasks. The StructBERT with structural pre-training gives surprisingly good empirical results on a variety of downstream tasks, including pushing the state-of-the-art on the GLUE benchmark to 89.0 (outperforming all published models), the F1 score on SQuAD v1.1 question answering to 93.0, the accuracy on SNLI to 91.7.		
	versions		versions			