

cases	doc_1		doc_2		decision	id
					DUPLICATES	252
			authors	<ul style="list-style-type: none">C. -C. Jay KuoBin Wang		
	authors	<ul style="list-style-type: none">Bin WangC.-C. Jay Kuo	title	SBERT-WK: A Sentence Embedding Method by Dissecting BERT-based Word Models		
	title	SBERT-WK: A Sentence Embedding Method by Dissecting BERT-Based Word Models	publication_date	2020-02-16 00:00:00		
	publication_date	2020-02-16 00:00:00	source	SupportedSources.PAPERS_WITH_CODE		
	source	SupportedSources.OPENALEX	journal			
	journal	IEEE/ACM transactions on audio, speech, and language processing	volume			
	volume	28	doi			
	doi	10.1109/taslp.2020.3008390	urls	<ul style="list-style-type: none">https://arxiv.org/pdf/2002.06652v2.pdfhttps://github.com/BinWang28/BERT_Sentence_Embedding		
	urls	<ul style="list-style-type: none">https://openalex.org/W3042631625https://doi.org/10.1109/taslp.2020.3008390http://arxiv.org/pdf/2002.06652	id	id-9106361643814454066		
	id	id-642621512081171246	abstract	Sentence embedding is an important research topic in natural language processing (NLP) since it can transfer knowledge to downstream tasks. Meanwhile, a contextualized word representation, called BERT, achieves the state-of-the-art performance in quite a few NLP tasks. Yet, it is an open problem to generate a high quality sentence representation from BERT-based word models. It was shown in previous study that different layers of BERT capture different linguistic properties. This allows us to fusion information across layers to find better sentence representation. In this work, we study the layer-wise pattern of the word representation of deep contextualized models. Then, we propose a new sentence embedding method by dissecting BERT-based word models through geometric analysis of the space spanned by the word representation. It is called the SBERT-WK method. No further training is required in SBERT-WK. We evaluate SBERT-WK on semantic textual similarity and downstream supervised tasks. Furthermore, ten sentence-level probing tasks are presented for detailed linguistic analysis. Experiments show that SBERT-WK achieves the state-of-the-art performance. Our codes are publicly available.		
	abstract		versions			
	versions					