cases	doc_1		doc_2		decision	id
	authors	• Bin Wang • C.-C. Jay Kuo	**authors**	• Bin Wang • C. -C. Jay Kuo	DUPLICATES	302
	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 19:02:52+00:00		
	source	SupportedSources.SEMANTIC_SCHOLAR	**source**	SupportedSources.ARXIV		
	journal	IEEE/ACM Transactions on Audio, Speech, and Language Processing	**journal**	None		
	volume	28	**volume**			
	doi	10.1109/taslp.2020.3008390	**doi**			
	urls	• https://www.semanticscholar.org/paper/d808e7468ef6265d39d3cd9c657c9f52e889cbc2	**urls**	• http://arxiv.org/pdf/2002.06652v2 • http://arxiv.org/abs/2002.06652v2 • http://arxiv.org/pdf/2002.06652v2		
	id	id-5252493220629724802	**id**	id7874431673068679021		
	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 fuse information across layers to find better sentence representations. 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**	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.		
	versions		**versions**			