		doc_1		doc_2	decision	id
cases	authors	• Timothee Mickus • Denis Paperno • Mathieu Constant • Kees van Deemter	authors	• Timothee Mickus • Denis Paperno • Mathieu Constant • Kees van Deemeter	DUPLICATES	51
	title	What do you mean, BERT? Assessing BERT as a Distributional Semantics Model	title	What do you mean, BERT? Assessing BERT as a Distributional Semantics Model		
	publication_date	2019-11-13 19:04:42+00:00	publication_date	2019-11-13 00:00:00		
	source	SupportedSources.ARXIV	source	SupportedSources.INTERNET_ARCHIVE		
	journal	Proceedings of the Society for Computation in Linguistics: Vol. 3 (2020), Article 34	journal			
	volume		volume			
	doi	10.7275/t778-ja71	doi			
	urls	• http://arxiv.org/pdf/1911.05758v2 • http://dx.doi.org/10.7275/t778-ja71 • http://arxiv.org/abs/1911.05758v2 • http://arxiv.org/pdf/1911.05758v2	urls	• https://web.archive.org/web/20200930174007/https://arxiv.org/pdf/1911.05758v1.pdf		
	id	id-460652827406954676	id	id-159479156452731012		
	abstract	Contextualized word embeddings, i.e. vector representations for words in context, are naturally seen as an extension of previous noncontextual distributional semantic models. In this work, we focus on BERT, a deep neural network that produces contextualized embeddings and has set the state-of-the-art in several semantic tasks, and study the semantic coherence of its embedding space. While showing a tendency towards coherence, BERT does not fully live up to the natural expectations for a semantic vector space. In particular, we find that the position of the sentence in which a word occurs, while having no meaning correlates, leaves a noticeable trace on the word embeddings and disturbs similarity relationships.	abstract	Contextualized word embeddings, i.e. vector representations for words in context, are naturally seen as an extension of previous noncontextual distributional semantic models. In this work, we focus on BERT, a deep neural network that produces contextualized embeddings and has set the state-of-the-art in several semantic tasks, and study the semantic coherence of its embedding space. While showing a tendency towards coherence, BERT does not fully live up to the natural expectations for a semantic vector space. In particular, we find that the position of the sentence in which a word occurs, while having no meaning correlates, leaves a noticeable trace on the word embeddings and disturbs similarity relationships.		
	versions		versions			