cases	doc_1		doc_2		decision	id
			authors	• Iryna Gurevych • Nils Reimers	DUPLICATES	327
	authors	• Nils Reimers • Iryna Gurevych	**title**	Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks		
	title	Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks	**publication_date**	2019-08-27 00:00:00		
	publication_date	2019-08-27 00:00:00	**source**	SupportedSources.PAPERS_WITH_CODE		
	source	SupportedSources.OPENALEX	**journal**			
	journal	arXiv (Cornell University)	**volume**			
	volume		**doi**			
	doi	10.48550/arxiv.1908.10084	**urls**	• https://arxiv.org/pdf/1908.10084v1.pdf • https://github.com/UKPLab/sentence-transformers • https://aclanthology.org/D19-1410.pdf		
	urls	• https://openalex.org/W4288253152 • https://doi.org/10.48550/arxiv.1908.10084 • http://arxiv.org/pdf/1908.10084	**id**	id8510547945187807396		
	id	id9010212408056506202	**abstract**	BERT (Devlin et al., 2018) and RoBERTa (Liu et al., 2019) has set a new state-of-the-art performance on sentence-pair regression tasks like semantic textual similarity (STS). However, it requires that both sentences are fed into the network, which causes a massive computational overhead: Finding the most similar pair in a collection of 10,000 sentences requires about 50 million inference computations (~65 hours) with BERT. The construction of BERT makes it unsuitable for semantic similarity search as well as for unsupervised tasks like clustering. In this publication, we present Sentence-BERT (SBERT), a modification of the pretrained BERT network that use siamese and triplet network structures to derive semantically meaningful sentence embeddings that can be compared using cosine-similarity. This reduces the effort for finding the most similar pair from 65 hours with BERT / RoBERTa to about 5 seconds with SBERT, while maintaining the accuracy from BERT. We evaluate SBERT and SRoBERTa on common STS tasks and transfer learning tasks, where it outperforms other state-of-the-art sentence embeddings methods.		
	abstract		**versions**			
	versions					