cases	doc_1		doc_2		decision	id
	authors	• AbÃlio M.P. De Jesus • Huilin Gao	**authors**	• Tonglei Guo • Huilin Gao	DUPLICATES	333
	title	Revisit Semantic Representation and Tree Search for Similar Question Retrieval.	**title**	Revisit Semantic Representation and Tree Search for Similar Question Retrieval		
	publication_date	2019-08-22 00:00:00	**publication_date**	2019-08-22 00:00:00		
	source	SupportedSources.OPENALEX	**source**	SupportedSources.SEMANTIC_SCHOLAR		
	journal	arXiv (Cornell University)	**journal**	ArXiv		
	volume		**volume**	abs/1908.08326		
	doi	None	**doi**			
	urls	• https://openalex.org/W2969493583	**urls**	• https://www.semanticscholar.org/paper/8ff3fc7e256fd47599ad00de132c30572b06c863		
	id	id468160454910591627	**id**	id824744131099170876		
	abstract		**abstract**	This paper studies the performances of BERT combined with tree structure in short sentence ranking task. In retrieval-based question answering system, we retrieve the most similar question of the query question by ranking all the questions in datasets. If we want to rank all the sentences by neural rankers, we need to score all the sentence pairs. However it consumes large amount of time. So we design a specific tree for searching and combine deep model to solve this problem. We fine-tune BERT on the training data to get semantic vector or sentence embeddings on the test data. We use all the sentence embeddings of test data to build our tree based on k-means and do beam search at predicting time when given a sentence as query. We do the experiments on the semantic textual similarity dataset, Quora Question Pairs, and process the dataset for sentence ranking. Experimental results show that our methods outperform the strong baseline. Our tree accelerate the predicting speed by 500%-1000% without losing too much ranking accuracy.		
	versions		**versions**			