cases	doc_1		doc_2		decision	id
			authors	• Mengjie Zhao • Philipp Dufter • Yadollah Yaghoobzadeh • Hinrich SchÃ¼tze	DUPLICATES	292
	authors	• Dufter, Philipp • SchÃ¼tze, Hinrich • Yaghoobzadeh, Yadollah • Zhao, Mengjie	**title**	Quantifying the Contextualization of Word Representations with Semantic Class Probing		
	title	Quantifying the Contextualization of Word Representations with Semantic Class Probing	**publication_date**	2020-04-25 17:49:37+00:00		
	publication_date	2020-01-01 00:00:00	**source**	SupportedSources.ARXIV		
	source	SupportedSources.CORE	**journal**	None		
	journal		**volume**			
	volume		**doi**			
	doi	10.5282/ubm/epub.74039	**urls**	• http://arxiv.org/pdf/2004.12198v2 • http://arxiv.org/abs/2004.12198v2 • http://arxiv.org/pdf/2004.12198v2		
	urls	• https://core.ac.uk/download/338940879.pdf	**id**	id-2352285641861899816		
	id	id-5836597132871540519	**abstract**	Pretrained language models have achieved a new state of the art on many NLP tasks, but there are still many open questions about how and why they work so well. We investigate the contextualization of words in BERT. We quantify the amount of contextualization, i.e., how well words are interpreted in context, by studying the extent to which semantic classes of a word can be inferred from its contextualized embeddings. Quantifying contextualization helps in understanding and utilizing pretrained language models. We show that top layer representations achieve high accuracy inferring semantic classes; that the strongest contextualization effects occur in the lower layers; that local context is mostly sufficient for semantic class inference; and that top layer representations are more task-specific after finetuning while lower layer representations are more transferable. Finetuning uncovers task related features, but pretrained knowledge is still largely preserved.		
	abstract	None	**versions**			
	versions					