	doc_1		doc_2		decision	id
cases	authors	• Hyeonbin Hwang • Haanju Yoo • Yera Choi	authors	• Choi, Yera • Hwang, Hyeonbin • Yoo, Haanju	DUPLICATES	169
	title	MED-SE: Medical Entity Definition-based Sentence Embedding	title	MED-SE: Medical Entity Definition-based Sentence Embedding		
	publication_date	2022-12-09 09:10:19+00:00	publication_date	2022-12-09 00:00:00		
	source	SupportedSources.ARXIV	source	SupportedSources.CORE		
	journal	None	journal			
	volume		volume			
	doi		doi	None		
	urls	• http://arxiv.org/pdf/2212.04734v1 • http://arxiv.org/abs/2212.04734v1 • http://arxiv.org/pdf/2212.04734v1	urls	• http://arxiv.org/abs/2212.04734		
	id	id6561121231106487290	id	id-3110811500391600908		
	abstract	We propose Medical Entity Definition-based Sentence Embedding (MED-SE), a novel unsupervised contrastive learning framework designed for clinical texts, which exploits the definitions of medical entities. To this end, we conduct an extensive analysis of multiple sentence embedding techniques in clinical semantic textual similarity (STS) settings. In the entity-centric setting that we have designed, MED-SE achieves significantly better performance, while the existing unsupervised methods including SimCSE show degraded performance. Our experiments elucidate the inherent discrepancies between the general- and clinical-domain texts, and suggest that entity-centric contrastive approaches may help bridge this gap and lead to a better representation of clinical sentences.	abstract	We propose Medical Entity Definition-based Sentence Embedding (MED-SE), a novel unsupervised contrastive learning framework designed for clinical texts, which exploits the definitions of medical entities. To this end, we conduct an extensive analysis of multiple sentence embedding techniques in clinical semantic textual similarity (STS) settings. In the entity-centric setting that we have designed, MED-SE achieves significantly better performance, while the existing unsupervised methods including SimCSE show degraded performance. Our experiments elucidate the inherent discrepancies between the general- and clinical-domain texts, and suggest that entity-centric contrastive approaches may help bridge this gap and lead to a better representation of clinical sentences.Comment: 8 pages, 2 figures, 9 table		
	versions		versions			