

cases	doc_1		doc_2		decision	id
	authors	<ul style="list-style-type: none">Sosuke NishikawaRyokan RiIkuya YamadaYoshimasa TsuruokaIsao Echizen	authors	<ul style="list-style-type: none">Sosuke NishikawaRyokan RiIkuya YamadaYoshimasa TsuruokaI. Echizen	DUPLICATES	166
	title	EASE: Entity-Aware Contrastive Learning of Sentence Embedding	title	EASE: Entity-Aware Contrastive Learning of Sentence Embedding		
	publication_date	2022-05-09 00:00:00	publication_date	2022-05-09 00:00:00		
	source	SupportedSources.INTERNET_ARCHIVE	source	SupportedSources.SEMANTIC_SCHOLAR		
	journal		journal			
	volume		volume			
	doi		doi	10.48550/arXiv.2205.04260		
	urls	<ul style="list-style-type: none">https://web.archive.org/web/20220511042006/https://arxiv.org/pdf/2205.04260v1.pdf	urls	<ul style="list-style-type: none">https://www.semanticscholar.org/paper/f5c39402fdf1df86cfd19256b44f7662aa357128		
	id	id-4472706160995810950	id	id-50210050938381095		
	abstract	We present EASE, a novel method for learning sentence embeddings via contrastive learning between sentences and their related entities. The advantage of using entity supervision is twofold: (1) entities have been shown to be a strong indicator of text semantics and thus should provide rich training signals for sentence embeddings; (2) entities are defined independently of languages and thus offer useful cross-lingual alignment supervision. We evaluate EASE against other unsupervised models both in monolingual and multilingual settings. We show that EASE exhibits competitive or better performance in English semantic textual similarity (STS) and short text clustering (STC) tasks and it significantly outperforms baseline methods in multilingual settings on a variety of tasks. Our source code, pre-trained models, and newly constructed multilingual STC dataset are available at https://github.com/studio-ousia/ease .	abstract	We present EASE, a novel method for learning sentence embeddings via contrastive learning between sentences and their related entities. The advantage of using entity supervision is twofold: (1) entities have been shown to be a strong indicator of text semantics and thus should provide rich training signals for sentence embeddings; (2) entities are defined independently of languages and thus offer useful cross-lingual alignment supervision. We evaluate EASE against other unsupervised models both in monolingual and multilingual settings. We show that EASE exhibits competitive or better performance in English semantic textual similarity (STS) and short text clustering (STC) tasks and it significantly outperforms baseline methods in multilingual settings on a variety of tasks. Our source code, pre-trained models, and newly constructed multi-lingual STC dataset are available at https://github.com/studio-ousia/ease .		
	versions		versions			