authors Column	 Xing Wu Chaochen Gao Yipeng Su Jizhong Han Zhongyuan Wang Songlin Hu Oothed Contrastive Learning for Unsupervised Sentence Embedding 22-09-12 00:00:00 OpportedSources.INTERNET_ARCHIVE	authors title publication_date source journal	Xing Wu Chaochen Gao Yipeng Su Jizhong Han Zhongyuan Wang Songlin Hu Smoothed Contrastive Learning for Unsupervised Sentence Embedding 2021-09-09 14:54:24+00:00 SupportedSources.ARXIV	
publication_date 2022-09 source Supporte journal volume doi id id-8668 Contrast previous (unsup-S similar s unsup-S	22-09-12 00:00:00	source	SupportedSources.ARXIV	
source Supporter journal volume doi id id-86687 Contrast previous (unsup-S similar s unsup-S				111
journal volume doi urls id id-8668* Contrast previous (unsup-S similar sunsup-S	oportedSources.INTERNET_ARCHIVE		None	(II
volume doi urls id id-8668' Contrast previous (unsup-5 similar sunsup-S		volume		ill
doi urls id id-86687 Contrast previous (unsup-8 similar sunsup-S		doi		ill
id id-8668 Contrast previous (unsup-8 similar sunsup-S			• http://arxiv.org/pdf/2109.04321v2	ill
Contrast previous (unsup-S similar s unsup-S	• https://web.archive.org/web/20220915033902/https://arxiv.org/pdf/2109.04321v2.pdf	urls	 http://arxiv.org/abs/2109.04321v2 http://arxiv.org/pdf/2109.04321v2 	DUPLICATES 1
previous (unsup-S similar s unsup-S	8668783657730631915	id	id2150374846178866255	ill
abstract size exculow-con smoothing InfoNCI negative unsup-S outperfo	ntrastive learning has been gradually applied to learn high-quality unsupervised sentence embedding. Among the vious un-supervised methods, the latest state-of-the-art method, as far as we know, is unsupervised SimCSE sup-SimCSE). Unsup-SimCSE uses the InfoNCE1loss function in the training stage by pulling semantically null resentences together and pushing apart dis-similar ones. Theoretically, we expect to use larger batches in sup-SimCSE to get more adequate comparisons among samples and avoid overfitting. However, increasing the ch size does not always lead to improvements, but instead even lead to performance degradation when the batch be exceeds a threshold. Through statistical observation, we find that this is probably due to the introduction of v-confidence negative pairs after in-creasing the batch size. To alleviate this problem, we introduce a simple cothing strategy upon the InfoNCE loss function, termedGaussian Smoothing InfoNCE (GS-oNCE). Specifically, we add random Gaussian noise vectors as negative samples, which act as a smoothing of the gative sample space. Though being simple, the proposed smooth-ing strategy brings substantial improvements to sup-SimCSE. We evaluate GS-InfoNCE on the standard semantic text similarity (STS)task. GS-InfoNCE performs the state-of-the-art unsup-SimCSE by an average Spear-man correlation of 1.38%, 0.72%, 1.17% 10.28% on the base of BERT-base, BERT-large, RoBERTa-base and RoBERTa-large, respectively.	abstract	Contrastive learning has been gradually applied to learn high-quality unsupervised sentence embedding. Among the previous un-supervised methods, the latest state-of-the-art method, as far as we know, is unsupervised SimCSE (unsup-SimCSE). Unsup-SimCSE uses the InfoNCE1loss function in the training stage by pulling semantically similar sentences together and pushing apart dis-similar ones. Theoretically, we expect to use larger batches in unsup-SimCSE to get more adequate comparisons among samples and avoid overfitting. However, increasing the batch size does not always lead to improvements, but instead even lead to performance degradation when the batch size exceeds a threshold. Through statistical observation, we find that this is probably due to the introduction of low-confidence negative pairs after in-creasing the batch size. To alleviate this problem, we introduce a simple smoothing strategy upon the InfoNCE loss function, termedGaussian Smoothing InfoNCE (GS-InfoNCE). Specifically, we add random Gaussian noise vectors as negative samples, which act as a smoothing of the negative sample space. Though being simple, the proposed smooth-ing strategy brings substantial improvements to unsup-SimCSE. We evaluate GS-InfoNCE on the standard semantic text similarity (STS)task. GS-InfoNCE outperforms the state-of-the-art unsup-SimCSE by an average Spear-man correlation of 1.38%, 0.72%, 1.17% and 0.28% on the base of BERT-base, BERT-large, RoBERTa-base and RoBERTa-large, respectively.	