cases	doc_1	doc_2	decision	id
	<table><tr><td>**authors**</td><td>• Haochen Tan • Wei Shao • Han Wu • Ke Yang • Linqi Song</td></tr><tr><td>**title**</td><td>A Sentence is Worth 128 Pseudo Tokens: A Semantic-Aware Contrastive Learning Framework for Sentence Embeddings</td></tr><tr><td>**publication_date**</td><td>2022-03-11 00:00:00</td></tr><tr><td>**source**</td><td>SupportedSources.OPENALEX</td></tr><tr><td>**journal**</td><td></td></tr><tr><td>**volume**</td><td></td></tr><tr><td>**doi**</td><td>None</td></tr><tr><td>**urls**</td><td>• https://openalex.org/W4226186836</td></tr><tr><td>**id**</td><td>id2628709532708716661</td></tr><tr><td>**abstract**</td><td></td></tr><tr><td>**versions**</td><td></td></tr></table>	<table><tr><td>**authors**</td><td>• Haochen Tan • Wei Shao • Han Wu • Ke Yang • Linqi Song</td></tr><tr><td>**title**</td><td>A Sentence is Worth 128 Pseudo Tokens: A Semantic-Aware Contrastive Learning Framework for Sentence Embeddings</td></tr><tr><td>**publication_date**</td><td>2022-03-11 00:00:00</td></tr><tr><td>**source**</td><td>SupportedSources.SEMANTIC_SCHOLAR</td></tr><tr><td>**journal**</td><td>ArXiv</td></tr><tr><td>**volume**</td><td>abs/2203.05877</td></tr><tr><td>**doi**</td><td>10.48550/arXiv.2203.05877</td></tr><tr><td>**urls**</td><td>• https://www.semanticscholar.org/paper/2886734184eb7efc1dca1b33c35d33bc41cdfe5c</td></tr><tr><td>**id**</td><td>id-5139767384257331520</td></tr><tr><td>**abstract**</td><td>Contrastive learning has shown great potential in unsupervised sentence embedding tasks, e.g., SimCSE (CITATION).However, these existing solutions are heavily affected by superficial features like the length of sentences or syntactic structures. In this paper, we propose a semantic-aware contrastive learning framework for sentence embeddings, termed Pseudo-Token BERT (PT-BERT), which is able to explore the pseudo-token space (i.e., latent semantic space) representation of a sentence while eliminating the impact of superficial features such as sentence length and syntax. Specifically, we introduce an additional pseudo token embedding layer independent of the BERT encoder to map each sentence into a sequence of pseudo tokens in a fixed length. Leveraging these pseudo sequences, we are able to construct same-length positive and negative pairs based on the attention mechanism to perform contrastive learning. In addition, we utilize both the gradient-updating and momentum-updating encoders to encode instances while dynamically maintaining an additional queue to store the representation of sentence embeddings, enhancing the encoderâ€™s learning performance for negative examples. Experiments show that our model outperforms the state-of-the-art baselines on six standard semantic textual similarity (STS) tasks. Furthermore, experiments on alignments and uniformity losses, as well as hard examples with different sentence lengths and syntax, consistently verify the effectiveness of our method.</td></tr><tr><td>**versions**</td><td></td></tr></table>	NOT DUPLICATES	383