<table>
<tr><th>cases</th><td>
<table>
<tr><th>doc_1</th><th>doc_2</th><th>decision</th><th>id</th></tr>
<tr><td>
<table>
<tr><th>authors</th><td>Tianyu Gao
Xingcheng Yao
Danqi Chen</td></tr>
<tr><th>title</th><td>SimCSE: Simple Contrastive Learning of Sentence Embeddings</td></tr>
<tr><th>publication_date</th><td>2022-05-18 00:00:00</td></tr>
<tr><th>source</th><td>SupportedSources.INTERNET_ARCHIVE</td></tr>
<tr><th>journal</th><td></td></tr>
<tr><th>volume</th><td></td></tr>
<tr><th>doi</th><td></td></tr>
<tr><th>urls</th><td>https://web.archive.org/web/20220526183956/https://arxiv.org/pdf/2104.08821v4.pdf</td></tr>
<tr><th>id</th><td>id9115064052159856573</td></tr>
<tr><th>abstract</th><td>This paper presents SimCSE, a simple contrastive learning framework that greatly advances state-of-the-art sentence embeddings. We first describe an unsupervised approach, which takes an input sentence and predicts itself in a contrastive objective, with only standard dropout used as noise. This simple method works surprisingly well, performing on par with previous supervised counterparts. We find that dropout acts as minimal data augmentation, and removing it leads to a representation collapse. Then, we propose a supervised approach, which incorporates annotated pairs from natural language inference datasets into our contrastive learning framework by using "entailment" pairs as positives and "contradiction" pairs as hard negatives. We evaluate SimCSE on standard semantic textual similarity (STS) tasks, and our unsupervised and supervised models using BERT base achieve an average of 76.3% and 81.6% Spearman's correlation respectively, a 4.2% and 2.2% improvement compared to the previous best results. We also show -- both theoretically and empirically -- that the contrastive learning objective regularizes pre-trained embeddings' anisotropic space to be more uniform, and it better aligns positive pairs when supervised signals are available.</td></tr>
<tr><th>versions</th><td></td></tr>
</table>
</td><td>
<table>
<tr><th>authors</th><td>Tianyu Gao
Xing-Cheng Yao
Danqi Chen</td></tr>
<tr><th>title</th><td>SimCSE: Simple Contrastive Learning of Sentence Embeddings</td></tr>
<tr><th>publication_date</th><td>2021-04-18 00:00:00</td></tr>
<tr><th>source</th><td>SupportedSources.OPENALEX</td></tr>
<tr><th>journal</th><td>arXiv (Cornell University)</td></tr>
<tr><th>volume</th><td></td></tr>
<tr><th>doi</th><td>10.48550/arxiv.2104.08821</td></tr>
<tr><th>urls</th><td>https://openalex.org/W3213189520
https://doi.org/10.48550/arxiv.2104.08821
http://arxiv.org/pdf/2104.08821</td></tr>
<tr><th>id</th><td>id-7788975266327500169</td></tr>
<tr><th>abstract</th><td></td></tr>
<tr><th>versions</th><td></td></tr>
</table>
</td><td>DUPLICATES</td><td>156</td></tr>
</table>
</td></tr>
</table>