| cases | | doc_1 | | doc_2 | decision | id |
|---|---|---|---|---|---|---|
| | authors | <ul><li>Poerner, Nina</li><li>SchÃ¼tze, Hinrich</li><li>Waltinger, Ulli</li></ul> | authors | <ul><li>Nina Poerner</li><li>Ulli Waltinger</li><li>Hinrich SchÃ¼tze</li></ul> | DUPLICATES | 310 |
| | title | Sentence Meta-Embeddings for Unsupervised Semantic Textual Similarity | title | Sentence Meta-Embeddings for Unsupervised Semantic Textual Similarity | | |
| | publication_date | 2020-01-01 00:00:00 | publication_date | 2019-11-09 14:31:03+00:00 | | |
| | source | SupportedSources.CORE | source | SupportedSources.ARXIV | | |
| | journal | | journal | None | | |
| | volume | | volume | | | |
| | doi | 10.5282/ubm/epub.72194 | doi | | | |
| | urls | <ul><li>https://core.ac.uk/download/323181766.pdf</li></ul> | urls | <ul><li>http://arxiv.org/pdf/1911.03700v3</li><li>http://arxiv.org/abs/1911.03700v3</li><li>http://arxiv.org/pdf/1911.03700v3</li></ul> | | |
| | id | id-6484347880971472127 | id | id-144447573610121757 | | |
| | abstract | We address the task of unsupervised Seman- tic Textual Similarity (STS) by ensembling di- verse pre-trained sentence encoders into sen- tence meta-embeddings. We apply, extend and evaluate different meta-embedding meth- ods from the word embedding literature at the sentence level, including dimensionality re- duction (Yin and Schu Ìˆtze, 2016), generalized Canonical Correlation Analysis (Rastogi et al., 2015) and cross-view auto-encoders (Bolle- gala and Bao, 2018). Our sentence meta- embeddings set a new unsupervised State of The Art (SoTA) on the STS Benchmark and on the STS12â€"STS16 datasets, with gains of be- tween 3.7% and 6.4% Pearsonâ€™s r over single- source systems | abstract | We address the task of unsupervised Semantic Textual Similarity (STS) by ensembling diverse pre-trained sentence encoders into sentence meta-embeddings. We apply, extend and evaluate different meta-embedding methods from the word embedding literature at the sentence level, including dimensionality reduction (Yin and Sch\"utze, 2016), generalized Canonical Correlation Analysis (Rastogi et al., 2015) and cross-view auto-encoders (Bollegala and Bao, 2018). Our sentence meta-embeddings set a new unsupervised State of The Art (SoTA) on the STS Benchmark and on the STS12-STS16 datasets, with gains of between 3.7% and 6.4% Pearson's r over single-source systems. | | |
| | versions | | versions | | | |