

cases	doc_1		doc_2		decision	id
	authors	<ul style="list-style-type: none">Chengwei WeiBin WangC. -C. Jay Kuo	authors	<ul style="list-style-type: none">Chen WeiBin WangC.-C. Jay Kuo	DUPLICATES	171
	title	SynWMD: Syntax-aware Word Mover's Distance for Sentence Similarity Evaluation	title	SynWMD: Syntax-aware Word Mover's Distance for Sentence Similarity Evaluation		
	publication_date	2022-06-20 22:30:07+00:00	publication_date	2022-06-20 00:00:00		
	source	SupportedSources.ARXIV	source	SupportedSources.SEMANTIC_SCHOLAR		
	journal	None	journal	ArXiv		
	volume		volume	abs/2206.10029		
	doi		doi	10.48550/arXiv.2206.10029		
	urls	<ul style="list-style-type: none">http://arxiv.org/pdf/2206.10029v1http://arxiv.org/abs/2206.10029v1http://arxiv.org/pdf/2206.10029v1	urls	<ul style="list-style-type: none">https://www.semanticscholar.org/paper/4ba496a286afe4a1efd196fedd25495be772ea60		
	id	id1041136267812084347	id	id735778374704121813		
	abstract	Word Mover's Distance (WMD) computes the distance between words and models text similarity with the moving cost between words in two text sequences. Yet, it does not offer good performance in sentence similarity evaluation since it does not incorporate word importance and fails to take inherent contextual and structural information in a sentence into account. An improved WMD method using the syntactic parse tree, called Syntax-aware Word Mover's Distance (SynWMD), is proposed to address these two shortcomings in this work. First, a weighted graph is built upon the word co-occurrence statistics extracted from the syntactic parse trees of sentences. The importance of each word is inferred from graph connectivities. Second, the local syntactic parsing structure of words is considered in computing the distance between words. To demonstrate the effectiveness of the proposed SynWMD, we conduct experiments on 6 textual semantic similarity (STS) datasets and 4 sentence classification datasets. Experimental results show that SynWMD achieves state-of-the-art performance on STS tasks. It also outperforms other WMD-based methods on sentence classification tasks.	abstract	Word Mover's Distance (WMD) computes the distance between words and models text similarity with the moving cost between words in two text sequences. Yet, it does not offer good performance in sentence similarity evaluation since it does not incorporate word importance and fails to take in-herent contextual and structural information in a sentence into account. An improved WMD method using the syntactic parse tree, called Syntax-aware Word Mover's Distance (SynWMD), is proposed to address these two shortcomings in this work. First, a weighted graph is built upon the word co-occurrence statistics extracted from the syntactic parse trees of sentences. The importance of each word is inferred from graph connectivities. Second, the local syntactic parsing structure of words is considered in computing the distance between words. To demonstrate the effectiveness of the proposed SynWMD, we conduct experiments on 6 textual semantic similarity (STS) datasets and 4 sentence classification datasets. Experimental results show that SynWMD achieves state-of-the-art performance on STS tasks. It also outperforms other WMD-based methods on sentence classification tasks.		
	versions		versions			