

cases	doc_1		doc_2		decision	id
			authors	<ul style="list-style-type: none"><li>L. Bajcetic</li><li>Seung-bin Yim</li></ul>	DUPLICATES	307
			title	Implementation of Supervised Training Approaches for Monolingual Word Sense Alignment: ACDH-CH System Description for the MWSA Shared Task at GlobaLex 2020		
			publication_date	2020-05-01 00:00:00		
			source	SupportedSources.SEMANTIC_SCHOLAR		
			journal			
			volume			
			doi			
			urls	<ul style="list-style-type: none"><li>https://www.semanticscholar.org/paper/00f54c753963f2ebde8ade314064b4737ef977d0</li></ul>		
			id	id-6429106363299514239		
			abstract	This paper describes our system for monolingual sense alignment across dictionaries. The task of monolingual word sense alignment is presented as a task of predicting the relationship between two senses. We will present two solutions, one based on supervised machine learning, and the other based on pre-trained neural network language model, specifically BERT. Our models perform competitively for binary classification, reporting high scores for almost all languages. This paper presents our submission for the shared task on monolingual word sense alignment across dictionaries as part of the GLOBALEX 2020 “ Linked Lexicography workshop at the 12th Language Resources and Evaluation Conference (LREC). Monolingual word sense alignment (MWSA) is the task of aligning word senses across resources in the same language. Lexical-semantic resources (LSR) such as dictionaries form valuable foundation of numerous natural language processing (NLP) tasks. Since they are created manually by experts, dictionaries can be considered among the resources of highest quality and importance. However, the existing LSRs in machine readable form are small in scope or missing altogether. Thus, it would be extremely beneficial if the existing lexical resources could be connected and expanded. Lexical resources display considerable variation in the number of word senses that lexicographers assign to a given entry in a dictionary. This is because the identification and differentiation of word senses is one of the harder tasks that lexicographers face. Hence, the task of combining dictionaries from different sources is difficult, especially for the case of mapping the senses of entries, which often differ significantly in granularity and coverage. (Ahmadi et al., 2020) There are three different angles from which the problem of word sense alignment can be addressed: approaches based on the similarity of textual descriptions of word senses, approaches based on structural properties of lexical-semantic resources, and a combination of both. (Matuschek, 2014) In this paper we focus on the similarity of textual descriptions. This is a common approach as the majority of previous work used some notion of similarity between senses, mostly gloss overlap or semantic relatedness based on glosses. This makes sense, as glosses are a prerequisite for humans to recognize the meaning of an encoded sense, and thus also an intuitive way of judging the similarity of senses. (Matuschek, 2014) The paper is structured as follows: we provide a brief overview of related work in Section 2, and a description of the corpus in Section 3. In Section 4 we explain all important aspects of our model implementation, while the results are presented in Section 5. Finally, we end the paper with the discussion in Section 6 and conclusion in Section 7.		
			versions			
		</				