cases	doc_1		doc_2		decision	id
	authors	• Alessandro Lenci • Magnus Sahlgren • Patrick Jeuniaux • Amaru Cuba Gyllensten • Martina Miliani	**authors**	• Alessandro Lenci • Magnus Sahlgren • Patrick Jeuniaux • Amaru Cuba Gyllensten • Martina Miliani	NOT DUPLICATES	403
	title	A comprehensive comparative evaluation and analysis of Distributional Semantic Models	**title**	A comparative evaluation and analysis of three generations of Distributional Semantic Models		
	publication_date	2021-05-20 00:00:00	**publication_date**	2021-05-20 15:18:06+00:00		
	source	SupportedSources.SEMANTIC_SCHOLAR	**source**	SupportedSources.ARXIV		
	journal	ArXiv	**journal**	None		
	volume	abs/2105.09825	**volume**			
	doi		**doi**			
	urls	• https://www.semanticscholar.org/paper/44ba2fade1d9369bb1643094a70be7052aae4fba	**urls**	• http://arxiv.org/pdf/2105.09825v2 • http://arxiv.org/abs/2105.09825v2 • http://arxiv.org/pdf/2105.09825v2		
	id	id-2126473246117565045	**id**	id-9075504213419007733		
	abstract	Distributional semantics has deeply changed in the last decades. First, predict models stole the thunder from traditional count ones, and more recently both of them were replaced in many NLP applications by contextualized vectors produced by Transformer neural language models. Although an extensive body of research has been devoted to Distributional Semantic Model (DSM) evaluation, we still lack a thorough comparison with respect to tested models, semantic tasks, and benchmark datasets. Moreover, previous work has mostly focused on task-driven evaluation, instead of exploring the differences between the way models represent the lexical semantic space. In this paper, we perform a comprehensive evaluation of type distributional vectors, either produced by static DSMs or obtained by averaging the contextualized vectors generated by BERT. First of all, we investigate the performance of embeddings in several semantic tasks, carrying out an in-depth statistical analysis to identify the major factors influencing the behavior of DSMs. The results show that i.) the alleged superiority of predict based models is more apparent than real, and surely not ubiquitous and ii.) static DSMs surpass contextualized representations in most out-of-context semantic tasks and datasets. Furthermore, we borrow from cognitive neuroscience the methodology of Representational Similarity Analysis (RSA) to inspect the semantic spaces generated by distributional models. RSA reveals important differences related to the frequency and part-of-speech of lexical items.	**abstract**	Distributional semantics has deeply changed in the last decades. First, predict models stole the thunder from traditional count ones, and more recently both of them were replaced in many NLP applications by contextualized vectors produced by Transformer neural language models. Although an extensive body of research has been devoted to Distributional Semantic Model (DSM) evaluation, we still lack a thorough comparison with respect to tested models, semantic tasks, and benchmark datasets. Moreover, previous work has mostly focused on task-driven evaluation, instead of exploring the differences between the way models represent the lexical semantic space. In this paper, we perform a comprehensive evaluation of type distributional vectors, either produced by static DSMs or obtained by averaging the contextualized vectors generated by BERT. First of all, we investigate the performance of embeddings in several semantic tasks, carrying out an in-depth statistical analysis to identify the major factors influencing the behavior of DSMs. The results show that i.) the alleged superiority of predict based models is more apparent than real, and surely not ubiquitous and ii.) static DSMs surpass contextualized representations in most out-of-context semantic tasks and datasets. Furthermore, we borrow from cognitive neuroscience the methodology of Representational Similarity Analysis (RSA) to inspect the semantic spaces generated by distributional models. RSA reveals important differences related to the frequency and part-of-speech of lexical items.		
	versions		**versions**			