cases	doc_1		doc_2		decision	id
	authors	• Iker GarcÃa-Ferrero • Rodrigo Agerri • German Rigau	**authors**	• GarcÃa Ferrero, Iker	DUPLICATES	308
	title	A Common Semantic Space for Monolingual and Cross-Lingual Meta-Embeddings	**title**	A common semantic space for monolingual and cross-lingual meta-embeddings		
	publication_date	2020-01-17 15:42:29+00:00	**publication_date**	2019-09-23 00:00:00		
	source	SupportedSources.ARXIV	**source**	SupportedSources.CORE		
	journal	None	**journal**			
	volume		**volume**			
	doi		**doi**	None		
	urls	• http://arxiv.org/pdf/2001.06381v2 • http://arxiv.org/abs/2001.06381v2 • http://arxiv.org/pdf/2001.06381v2	**urls**	• https://core.ac.uk/download/547392015.pdf		
	id	id6441832917161128703	**id**	id55403628157404018l0		
	abstract	This paper presents a new technique for creating monolingual and cross-lingual meta-embeddings. Our method integrates multiple word embeddings created from complementary techniques, textual sources, knowledge bases and languages. Existing word vectors are projected to a common semantic space using linear transformations and averaging. With our method the resulting meta-embeddings maintain the dimensionality of the original embeddings without losing information while dealing with the out-of-vocabulary problem. An extensive empirical evaluation demonstrates the effectiveness of our technique with respect to previous work on various intrinsic and extrinsic multilingual evaluations, obtaining competitive results for Semantic Textual Similarity and state-of-the-art performance for word similarity and POS tagging (English and Spanish). The resulting cross-lingual meta-embeddings also exhibit excellent cross-lingual transfer learning capabilities. In other words, we can leverage pre-trained source embeddings from a resource-rich language in order to improve the word representations for under-resourced languages.	**abstract**	This masterâ€™s thesis presents a new technique for creating monolingual and cross-lingual meta-embeddings. Our method integrates multiple word embeddings created from complementary techniques, textual sources, knowledge bases and languages. Existing word vectors are projected to a common semantic space using linear transformations and averaging. With our method the resulting meta-embeddings maintain the dimensionality of the original embeddings without losing information while dealing with the out-of-vocabulary (OOV) problem. Furthermore, empirical evaluation demonstrates the effectiveness of our technique with respect to previous work on various intrinsic and extrinsic multilingual evaluations		
	versions		**versions**			