

language	dataset	parent dataset	Skipgram	fastText, h=10, bigram	MDLText
		average	0,00%	94,65%	84,57%
EN	20Newsgroups	-			90,20%
MULTI	Fbis	TREC			79,00%
EN	Irish economic	-			67,20%
EN	La1S	TREC-5			87,70%
EN	La2S	TREC-5			88,30%
EN	Latimes	LA Times			84,70%
EN	New3S	-			85,80%
EN	RCV1 Top-4	-			92,00%
IT	Rcv2-Italian	-			79,40%
PT	Rcv2-Portuguese	-			79,10%
SP	Rcv2-Spanish	-			84,10%
EN	Re0	Reuters-21578			81,50%
EN	Re1	Reuters-21578			81,20%
EN	Re8	Reuters-21578			91,40%
EN	Reuters	-			91,20%
EN	Reviews	San Jose Mercury newspaper			90,30%
EN	AG News	-		92,50%	
ZN	Sogou News	-		96,80%	