



Figure 1: Performance change as the corpus size growing (a) on the WS353 word similarity task; (b) on the Google word analogy task on English and Czech datasets. Results show that PU-Learning model consistently outperforms SGNS and GloVe when the size of corpus is small.

Language	Similarity task					Analogy task
	WS353	Similarity	Relatedness	M. Turk	MEN	Google
English (en)	353	203	252	287	3000	19544
Czech (cs)	337	193	241	268	2810	18650
Danish (da)	346	198	247	283	2951	18340
Dutch (nl)	346	200	247	279	2852	17684

Table 1: The size of test sets in all languages. English test sets are the original test sets. Test sets in other languages are translated from English test sets