1 256 256 1.0 0.0005 30 128		dim	$dim\_word$	lr	decay- $c$	maxlen	batch sie
2 252 252 10 0001 50 00	1	256	256	1.0	0.0005	30	128
$2 \mid 256  256  1.0  0.001  50  32$	2	256	256	1.0	0.001	50	32