

Task	Dataset	Level	Vocab	AvgLen	#train	#valid	#test
Number Prediction	synthetic	word	100	100 words	1M	10K	10K
Sentiment Analysis	Rotten Tomatoes	word	18,764	22 words	8,835	1,079	1,079
Sentiment Analysis	IMDB	word	112,540	241 words	21,143	3,857	25,000
News Classification	AG	character	70	200 characters	101,851	18,149	7,600
Q/A	Children Book Test-NE	sentence	53,063	20 sentences	108,719	2,000	2,500
Q/A	Children Book Test-CN	sentence	53,185	20 sentences	120,769	2,000	2,500