Words\_219k.txt为原始语料库

words\_frequency.txt为原始语料库的词频排布

Word0-19为按频率划分梯度的单词集合