

1. ngram-count -text train.txt -order 3 -write train.count

-text 指向输入文件

-order 指向生成几元的 n-gram, 即 n

-write 指向输出文件

2. ngram-count -read train.count -order 3 -lm model.lm -interpolate -kn-counts-modified

-read 指向输入文件，为上一步的输出文件

-order 与上同

-lm 指向训练好的语言模型输出文件

Kneser-Ney smoothing-**kn-counts-modified**

3. ngram -ppl test.txt -order 3 -lm model.lm > result

-ppl 为对测试集句子进行评分( $\log P(T)$ ，其中  $P(T)$  为所有句子的概率乘积) 和计算测试集困惑度的参数  
result 为输出结果文件