1\_process.py：

python 1\_process.py zhwiki-latest-pages-articles.xml.bz2 wiki.zh.txt

opencc:进入解压后的opencc的目录，双击opencc.exe文件，在当前目录打开dos窗口，输入如下命令

opencc -i E:\wiki\_zh\_word2vec-master\wiki.zh.txt -o E:\wiki\_zh\_word2vec-master\wiki.zh.simp.seg.txt -c t2s.json

3\_train\_word2vec\_model.py：

python 3\_train\_word2vec\_model.py wiki.zh.simp.seg.txt wiki.zh.model wiki.zh.vector

备注：

wiki.zh.txt为wiki中文语料源文件

wiki.zh.simp.txt为wiki中文语料繁体转换成简体后的文件

wiki.zh.simp.seg.txt为wiki中文语料简体分词后的结果