1、数据处理：

数据集为'dataset/news.txt'，先进行数据预处理得到分词后的文件：

2、训练word2vec

得到cutdata.txt，再利用word2vec模型进行词嵌入：

3、测试word2vec

得到word\_embedding.txt文件，使用“相似度”测试效果。

示例如下：输入单词“男人”，得到词表中与“男人”最相关的单词

