# similar sentences gen v1.1 使用说明

调用 **generate\_sentences.py** 中的 generate\_similar\_sentences 函数。

两种生成方式：同义词替换和回译法，默认情况下都使用

生成方式选择：修改 **generate\_sentences.py** 中第 72 行对 generate 函数的调用参数

generate(original\_sentence, custom\_dict\_file, similar\_dict\_file, synonymous=True, trans=False) 即表示使用同义词替换而不使用回译法。

synonymous, trans 两个参数默认为 True

具体的替换实现在 segmentation.py 中，在分词的基础上进行了修改。

**输入：**

原句txt文件

**输出:**

生成一个 output\_file 文件, 第一列为原句，第二列为生成的相似句，第三列为标签。

**函数接口：**

def generate\_similar\_sentences(  
 input\_file,

output\_file,  
 column,  
 label,  
 custom\_dict\_file,  
 similar\_dict\_file,  
)

**参数说明：**

:param input\_file: 要生成同义句的句子文件

:param output\_file: 输出文件  
:param column: 原句在文件中的列号  
:param label: 生成句子的标签  
:param similar\_dict\_file: 同义词字典文件  
:param custom\_dict\_file: 自定义字典文件

使用：

修改 main 函数中的 input\_file 和 output\_file，运行 generate\_sentences.py