## 1.Word\_edit\_distance

1.

Matrix = np.zeros((len2,len1))

注意：len2表示垂直方向，len1表示水平方向，类似于矩阵3✖2

0 0

0 0

0 0

2.range（0，len(x)）:

len(x)不被包括，0被包括

3.

Python中对于数字的不等于 ！=

对于否定，不是！，而是not

例如：if not function（）：

4.列表反转

list.reverse()

## 2.Sentence\_edit\_distance

5.句子变换为列表split()

str1 = x.split()

6.

列表插入元素

x.insert(index,element)

列表末尾加入元素

x.append(element)

## 3.word\_batch()

7.文件读取

**注意同一个文件下的文件名不需要添加.txt**

**后面加encoding = “UTF-8” 可以除去txt文件中有例如中文的麻烦**

with open(inputfile,‘r‘，encoding = “UTF-8”) as input:

with open（outputfile，“a”, encoding = UTF-8”）as output:

lines = f.readlines();#整行读取并且按空格化作表格

for line in lines:

word = line.split();

8.rstrip()

注意文件读取时候，例如上面的word[-1]不仅包含最后一个字符，还有很多空白也包含，这会严重影响输出结果

因此为了只得到字符，需要word.rstrip()