

第六次作业内容

0 / 1

编程题

1.# 请注释掉下面代码中的pass，填充相应的代码

函数输入：文件路径（见调用示例）
函数输出：该路径下所有文件中单词出现的频率（以字典数据类型进行返回，去除掉长度不大于2的单词）
注意：1、在统计单词出现的频率时，所有的单词均应先变成小写。
2、在统计单词时，从非字母字符（包括空格）的下一个字符作为单词的字母，直到
非字母字符（包括空格）截止。
3、数据为整个email文件夹
例如： Increase volume of Ejacu1ate 分离出来的单词为increase ,volume , ofejacu
,ate

```
def fact(path):  
# 在下方填充代码，注释掉pass  
pass
```

```
# 主函数  
if __name__ == "__main__":  
# 示例:  
print(fact('./email/'))
```

你的答案：

```
import os  
import re  
  
def fact(path):  
    word_frequency = {} # 用于存储单词频率的字典  
  
    # 遍历指定路径下的所有文件  
    for root, dirs, files in os.walk(path):  
        for file in files:  
            file_path = os.path.join(root, file)  
            # 使用指定编码打开文件  
            with open(file_path, 'r', encoding='ISO-8859-15') as f:  
                content = f.read()  
                # 将内容转换为小写字母
```

[返回](#)