基于用户过滤推荐

1、问题描述

电影	张三	李四	王五	我
《新白娘子传奇》	1	4	3	5
《笑傲江湖》	1	0	4	5
《龙门飞甲》	0	2	1	0
《泰坦尼克号》	1	1	4	5
《流浪地球》	1	1	4	值得看么？

2、解决方案

```
import pandas as pd
from math import sqrt

def load_data(file):
    data = pd.read_csv(file)
    return data


def get_rating(data, index):
    '''
    返回第index这个人的电影的评分，返回类型是[]
    '''
    result =[]
    result = data[data["user_id"]==index]["rating"].values
    return  result

def computer_distance(rating_index, rating_me):
    #x[1,1,0,1]
    #y[5,5,0,5]
    sum = 0
    for i in range(0, 4):
        sum = sum+ (rating_index[i]-rating_me[i])*(rating_index[i]-rating_me[i])
    return sqrt(sum)

def get_distance_with_me(data, index):
    '''
    获取第index人与我的距离
    '''
    # 1、获取index这个人的数据（电影的评分） [1, 2,,4,5]
```

```
    rating_index = []
    rating_me = []
    rating_index = get_rating(data, index)
    print(rating_index)
    # 2、获取我这个人的数据（电影的评分）[1,2,3,4]
    rating_me = get_rating(data, 4)
    # 3、利用欧式距离公式计算
    result = computer_distance(rating_index, rating_me)
    return result


def find_min_distance(map):
    index = 0
    distance = 0
    min = 99999
    for key, value in map.items():
        print("key: %s , vlalue:%s" %(key, value))
        if min > value:
            min = value
            index = key

    distance = min
    return index, distance
```

```
# 1、数据到导入，给我这个程序使用
file = "E:\\movie.csv"
data = load_data(file)
print(data[data["user_id"]==1]["rating"].values)

# 2、计算每一个人和我的距离
# distance_1 = get_distance_with_me(data, 1)
# distance_2 = get_distance_with_me(data, 2)
# distance_3 = get_distance_with_me(data, 3)
map = {}
for i in range(1, 4):
    map[i] = get_distance_with_me(data, i) #{1:xx, 2:xx, 3:xx} i:表示第i个人
    print(map[i])

# # 3、找出谁与我距离最短，
index, distance = find_min_distance(map) #{1:xx, 2:xx, 3:xx}
print("第%s人跟我的距离最接近，他的值是%s" %(index, distance))

# 4、利用这个“谁”给我推荐
```

```
[1 1 0 1]
[1 1 0 1]
6.928203230275509
[4 0 2 1]
6.782329983125268
[3 4 1 4]
2.6457513110645907
key: 1 , vlalue:6.928203230275509
key: 2 , vlalue:6.782329983125268
key: 3 , vlalue:2.6457513110645907
第3人跟我的距离最接近，他的值是2.6457513110645907
```

```python
#1+2+3+4...10
sum = 0
for i in range(1, 11):
    sum =  sum+i
print(sum)
```

```
55
```

```
```