

基于用户过滤推荐

一、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？

二、解决方案

1、找出相近的用户

```
import pandas as pd
from math import sqrt
```

```
def get_person_rating(data):
    person_rating= {}
    # {1:[1,1,0,1],2:[4,0,2,1]} 1第一个用户 [1,1,0,1]
    # 对第一部电影
    for i in range(1, 5):
        person_i = data[data["user_id"]==i]["rating"].values
        person_rating[i] = person_i

    print(person_rating)
    return person_rating

def distance(person_rating, index) :
    it = person_rating[index]
    me = person_rating[4]

    sum =0
    for i in range(0, 4):
        sum = sum + pow((it[i]-me[i]),2)

    return sqrt(sum)
```

```
# 1、导入数据给这个程序使用
data = pd.read_csv("C:\\Users\\Administrator\\Desktop\\movie.csv")
#print(data)

#print(data[data["user_id"]==1]["rating"])
#print(data[data["user_id"]==2]["rating"])
# 2、 获取每一个人对电影的评分，要求{1:[1,1,0,1],2:[4,0,2,1]}
person_rating = {}
person_rating = get_person_rating(data)

# 3、其他3个用户我的距离
min = 6
min_index = 0
for i in range(1, 4):
    distance_person_i = distance(person_rating, i)
    if distance_person_i < min:
        min = distance_person_i
        min_index = i
print(min)
print(min_index)
```

```
{1: array([1, 1, 0, 1], dtype=int64), 2: array([4, 0, 2, 1], dtype=int64), 3: array([3, 4,
1, 4], dtype=int64), 4: array([5, 5, 0, 5], dtype=int64)}
2.6457513110645907
3
```

2、给我推荐电影