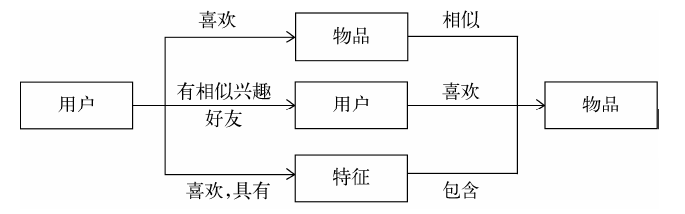
一、推荐系统的主要方法

推荐系统的目的是联系用户的兴趣和物品，这种联系需要依赖不同的媒介。目前流行的推荐系统基本上通过以下3种方式联系用户兴趣和物品：

（一）利用用户喜欢过的物品，给用户推荐与他喜欢过的物品相似的物品。

（二）利用和用户兴趣相似的其他用户，给用户推荐那些和他们兴趣爱好相似的其它用户喜欢的物品。

（三）通过一些特征联系用户和物品，给用户推荐那些具有用户喜欢的特征的物品。这里的特征有不同的表现方式，可以是物品的属性集合，也可以是隐语义向量。



二、测试数据

为了便于理解，这里只很了很少的数据量。tagdata.txt中的数据为：

```

刘一 黄飞鸿 movie

刘一 风清扬 music

陈二 黄飞鸿 movie

陈二 黄昏 music

张三 黄飞鸿 movie

张三 新龙门客栈 movie

张三 爱就一个字 music

李四 新龙门客栈 movie

李四 爱就一个字 music

李四 古拉格群岛 book

王五 爱就一个字 music

```

三、对应关系

从测试数据，可以计算出：

（一）用户-物品对应关系

|  |  |  |
| --- | --- | --- |
| 用户 | 物品 | 物品数量 |
| 刘一 | 黄飞鸿 | 1 |
| 风清扬 | 1 |
| 陈二 | 黄飞鸿 | 1 |
| 黄昏 | 1 |
| 张三 | 黄飞鸿 | 1 |
| 新龙门客栈 | 1 |
| 爱就一个字 | 1 |
| 李四 | 新龙门普天率土 | 1 |
| 爱就一个字 | 1 |
| 古拉格群岛 | 1 |
| 王五 | 爱就一个字 | 1 |

（二）用户-标签对应关系

|  |  |  |
| --- | --- | --- |
| 用户 | 标签 | 标签数量 |
| 刘一 | movie | 1 |
| music | 1 |
| 陈二 | movie | 1 |
| music | 1 |
| 张三 | movie | 2 |
| music | 1 |
| 李四 | movie | 1 |
| music | 1 |
| book | 1 |
| 王五 | music | 1 |

（三）物品-标签对应关系

|  |  |  |
| --- | --- | --- |
| 物品 | 标签 | 标签数量 |
| 黄飞鸿 | movie | 3 |
| 风清扬 | music | 1 |
| 黄昏 | music | 1 |
| 新龙门客栈 | movie | 2 |
| 爱就一个字 | music | 3 |
| 古拉格群岛 | book | 1 |

（四）标签-物品对应关系

|  |  |  |
| --- | --- | --- |
| 标签 | 物品 | 物品数量 |
| movie | 黄飞鸿 | 3 |
| 新龙门客栈 | 2 |
| music | 风清扬 | 1 |
| 黄昏 | 1 |
| 爱就一个字 | 3 |
| book | 古拉格群岛 | 1 |

三、标签流行度

四、余弦相似度

物品之间的相似度可以用物品标签向量的余弦相似度来衡量。

五、推荐列表多样性

六、程序

```

```

运行结果