**L1-exercise – Qiang Alfred Chen**

**Thinking-1 既然内容相似度计算简单，能频繁更新，为什么还需要协同过滤算法呢？**

内容过滤推荐会基于与已选物品的相似度来进行推荐。主要步骤有：1）通过在抓取每个商品的一系列特征来构建物品档案；2）通过用户购买的物品特征来构建基于内容的用户档案；3）通过特定的相似度方程计算用户档案和物品档案的相似度；4）推荐相似度最高的若干个物品。

所以它会一直推荐和用户已选物品相类似的物品，而不会推荐和已选物品差别很大的物品。而用户的需求其实是非常多样化的。

针对这一点，协同过滤一般采用最近邻技术，利用用户的历史喜好信息计算用户之间的距离，然后利用目标用户的最近邻居用户对商品评价的加权评价值来预测目标用户对特定商品的喜好程度，系统从而根据这一喜好程度来对目标用户进行推荐，满足用户非常多样化的潜在推荐需求。而且，和内容过滤相比，协同过滤对推荐对象没有特殊的要求，能处理非结构化的复杂对象，如音乐、电影。

**Thinking2 你需要推荐系统么？哪些情况下不需要推荐系统？**

随着网上信息量的大幅增长，用户在面对大量信息时无法从中获得对自己真正有用的那部分信息，对信息的使用效率反而降低了，出现[信息超载](https://baike.baidu.com/item/%E4%BF%A1%E6%81%AF%E8%B6%85%E8%BD%BD)（Information Overload）问题。[推荐系统](https://baike.baidu.com/item/%E6%8E%A8%E8%8D%90%E7%B3%BB%E7%BB%9F)是这个问题非常有潜力的方案之一，它是根据用户的信息需求、兴趣等，将用户感兴趣的信息、产品等推荐给用户。

所以出现[信息超载](https://baike.baidu.com/item/%E4%BF%A1%E6%81%AF%E8%B6%85%E8%BD%BD)时，[推荐系统](https://baike.baidu.com/item/%E6%8E%A8%E8%8D%90%E7%B3%BB%E7%BB%9F)是非常有用的工具。

但[信息不超载](https://baike.baidu.com/item/%E4%BF%A1%E6%81%AF%E8%B6%85%E8%BD%BD)时，[推荐系统](https://baike.baidu.com/item/%E6%8E%A8%E8%8D%90%E7%B3%BB%E7%BB%9F)的实际效用不明显，[推荐系统](https://baike.baidu.com/item/%E6%8E%A8%E8%8D%90%E7%B3%BB%E7%BB%9F)的运营成本反而成了额外的负担。

**Thinking3 如果给一个视频打标签，视频中有音乐作为背景音乐，采用了NLP方式对内容自动打标签，可能存在什么问题？**

视频是非结构化的复杂对象，视频中的背景音乐和视频的关联常常是非结构化的。这使得采用了NLP方式对内容自动打标签，标签的准确率会很不理想。