****

**学士学位本科生**

**开题报告与文献总结**

题目：基于Android的电子书推荐系统的设计与实现

|  |  |
| --- | --- |
| 学生姓名： | 刘亚辉 |
| 学 号： | 1907010317 |
| 专业班级： | 计算机科学与技术19-3班 |
| 指导教师： | 张红霞 |

2023年3月1日

**开题报告**

# 1 选题目的及研究意义

互联网的高速发展，改变了我们的生活方式，打破了我们的时空界限，重塑着我们的社会形态。在这个大数据的时代，我们的生活中随处可见智能推荐系统，它可以向用户推荐用户想要看到的内容，便于用户在庞大的信息中，准确找到有用的内容。推荐系统一方面为用户的使用提供了便捷，另一方面还可以为系统增加用户粘性。本文的选题是，该研究的目的是设计出一种能够个性化、精准地为用户推荐电子书的智能化推荐系统。该系统的设计需要考虑用户的兴趣、阅读习惯、历史阅读记录等多方面因素，通过机器学习、大数据分析等技术，为用户提供更优质、个性化的电子书推荐服务，从而提高用户的阅读体验和满意度 。

研究意义可以简单分为下面几个方面：

1.提高阅读体验：通过推荐系统，用户可以更快速、准确地找到自己感兴趣的电子书，从而提高阅读体验和效率

2.促进电子书市场发展：电子书市场的发展需要更多的用户和更高的销售量，而推荐系统可以帮助用户更好地发现和购买电子书，从而促进电子书市场的发展。

3.推动智能化技术发展：电子书智能推荐系统需要借助人工智能、大数据等技术，因此研究该系统也有助于推动智能化技术的发展和应用。

4.为图书馆和出版社提供参考：电子书智能推荐系统的研究成果可以为图书馆和出版社提供参考，帮助他们更好地了解用户需求和市场趋势，从而更好地为读者提供服务。

# 2 国内外研究现状

目前，电子书智能推荐系统的研究在国内外都有较为广泛的应用和研究。国内外学者都对基于用户兴趣和行为的推荐算法进行了深入研究，如基于协同过滤、基于深度学习、基于知识图谱、基于内容过滤和基于混合过滤等算法。[1]取得了不错的成果，并且该研究方向已经有部分内容应用在了互联网平台上，比如国内的京东、当当、豆瓣等，国外的Amazon、Google Books、The New York Public Library等。

## 2.1 国外研究现状

个性化推荐的概念是在1995年美国人工智能学会上提出的，并在以后的时间得到蓬勃发展并应用到各个领域之中[2]。

在电商领域，被誉为“推荐系统之王”的Amazon是美国最大的电商网站，据统计，Amazon网站上对用户的个性化推荐贡献了页面30%浏览量。很多个性化推荐方面的研究也是由 Amazon主导的。1998年 Amazon希望根据用户过去的消费和评分记录来为用户产生个性化推荐，由此提出了一种基于商品的协同过滤推荐（Item-CF）算法[3]。这一算法因为结合了用户的行为特征，可以显著地提高推荐结果的精确度，自提出之后，便被广泛传播和应用。在社交领域，国外的典型应用代表有Facebook和Twitter，他们在为用户推荐好友和内容的时候也会根据用户过去的评论、点赞等行为信息，从而找到用户偏好的内容。在广告展示方面，雅虎是其中的佼佼者。为了能够更加精准地投放广告，它根据用户对所投放的广告的反馈，利用LFM（隐语义)模型进行了网站的个性化首页设计。而电视产品个性化推荐方面最有名的当属Netflix——一个著名的在线视频网站。2009年Netflix 发起的Netflix Prize百万美金竞赛，积极引导了专业人员对个性化推荐的研究。随后几年 Netflix先后提出Personalized Video Ranker、Top-N VideoRanker、Trending Now、Continue Watching、Page Generation等核心推荐算法[4]，并取得商业上的巨大成功。

## 2.2 国内研究现状

在国内，电子商务、视频播放、社交网络等多个平台也在广泛使用和改进个性化推荐技术。但学者们慢慢注意到主流的个性化推荐算法虽然被大众所采纳和接受，但仍然存在着冷启动、推荐结果不够个性化等问题。为此，学者们进行了多种研究，或是对推荐算法本身加以改进，又或者是将多种算法进行组合形成新的推荐算法。在算法的相似度计算方面，陈功平等[5]引入了惩罚系数来改进传统的皮尔逊相关系数的计算方法;H.Sun等认为在计算相似度时，要考虑用户评价项目涉及的范围，因此他提出了一种杰卡德统一算子距离的计算方法，引入Jaccard系数来衡量两个用户的评分项目的重叠程度;在特征的扩展方面，高长元等充分利用用户对项目属性的偏好，扩展了用户偏好模型;在混合算法方面，，王婵在传统的协同过滤推荐中加入了权重,因子，使用权重组合来进行算法混合;王睿等[6]混合了项目的横向和纵向相似度，将不同层次的相似度进行配对结合形成混合推荐,改善了推荐精度。

在针对电视节目的个性化推荐上面，由于很难获取用户对电视节目的直接评分来构建用户评分矩阵，学者们逐渐意识到在个性化推荐方面，应该注重内容特征和文字信息的情感倾向，因此基于上下文信息进行推荐的方法逐渐流行。刘晓琳[7]认为没有单一存在的项目，她注重项目间的关联关系，在基于用户的协同过滤推荐上，为项目抽取了关联规则，混合了基于关联规则的推荐算法，改善了单一算法的不足;吴晓亮[8]认为，用户的评论会影响之后的用户观看倾向，因此他根据用户评论中含有的情感倾向，提出了基于用户评论情感的的矩阵分解个性化推荐。

可以看出，个性化推荐技术虽然发展已久，但对推荐算法存在问题的研究一直在继续。没有一个算法是十全十美的，都有自己的适用场景，因此在对个性化推荐算法进行改进时需要因势利导。

# 3 相关技术原理及技术路线

## 3.1 技术原理

1.Android客户端

开发平台：基于Android平台进行开发。开放源代码操作系统（OS）的典型代表就是 Android, 伴随着安卓迅速发展，基于Linux为核心的Android手机操作平台，因为其开放特点，程序的开发人员目前可以采取多种语言进行安卓应用软件研发。开发者针对 Android，能够采取SL4A来使用python、lua、tcl、php等各种脚本语言进行编程，能够应用编程语言Java等研发应用APP软件。而在网页编程语言方面,ajax,JavaScript,HTML5,在安卓应用软件中可以使用还包括dojo ,jquery其他多种语言。

UI显示：用户登录后客户端调用Web Service特定的方法接口服务端返回JSON格式的推荐结果客户端使用List View组件显示个性化的推荐图书列表。客户端采用异步方式获取数据提高用户体验避免推荐算法的运算时间过长所造成的假死现象。用户点击感兴钮的图书,得到图书详情页面而后再设置图书标签和评分这些显式信息即是服务端向客户说提供准确推荐结果的重要依据通过Web Service接口可实时上传至服务端。

网络请求：使用OKHttp3技术进行网络请求，OKHttp是一个当前主流的网络请求的开源框架，它支持GET和POST请求，支持Http的文件上传和下载，支持加载图片，支持下载文件透明的GZIP压缩，支持响应缓存避免重复的网络请求，支持使用连接池来降低响应延迟的问题。

2.服务器端：

使用Python Flask搭建一个轻量级服务器。 Flask是一个轻量级的可定制框架，使用Python语言编写，较其他同类型框架更为灵活、轻便、安全且容易上手。Flask的基本模式为在程序里将一个视图函数分配给一个URL，每当用户访问这个URL时，系统就会执行给该URL分配好的视图函数，获取函数的返回值。

数据库：数据库是用来管理组织数据的台库，随着现代网络和信息技术的发展，数据库已转变成用户所迫切需要的数据管理工具。其中， MySQL数据库是一款高安全性、高效率且可跨平台的数据库系统，同时可与PHP，Java等主流编程语言结合使用，具有很强的兼容性。MySQL是一种关系型数据库管理系统，关系数据库将数据保存在不同的表中，而不是将所有数据放在一个大仓库内，这样就增加了速度并提高了灵活性。本文推荐系统的数据存储在MySQL数据库中，由于其具有安全性高及存储容量大的特点，因此成为支撑推荐系统数据的不二选择。

3.推荐算法

基于内存的协同过滤推荐技术，其主要由三分部构成:用户对项目偏好的表示，相似近邻选取，为目标用户产生推荐。[9]

3.1用户对项目偏好表示

在用户对项目偏好表示这一步中，可以根据用户在系统中的历史行为数据得到， 在用户项目偏好表示方面，相对有效而又简单地做法是直接利用用户-项目评分矩阵中用户对项目的评分来表示用户对项目的偏好。而矩阵中的评分值则代表用户对项目的偏好程度。矩阵中的评分的取值范围一般为1-5，如果用户没有对项目进行评分则值为0，而评分从1到5则代表用户对项目的偏好程度逐渐变高。假设系统中有m个用户构成一个用户集合，n个项目构成一个项目集合，则用户对项目的评分用用户-项目（User-Item）评分矩阵表示：，

其中，元素为用户u对项目i的评分。

3.2相似近邻选取

相似近邻的选取最为关键，其直接关系到最后的推荐质量，如果相似近邻选取的过多则有可能因为多数邻居与目标用户的相似度不高而降低推荐的准确率，另一方面如果近邻过少则有可能因为近邻购买的产品用户也都购买过而无法产生新颖的推荐。而根据是基于用户去计算相似度还是基于项目去计算相似度又可以将基于内存的协同过滤推荐分为:基于用户的（user-based）协同过滤和基于项目(item-based)的协同过滤。[9]下面是传统的相似度计算方法：

皮尔森相关性系数（Pearson Correlation Coefficient），用来反映两个变量线性相关程度。其运用于推荐系统中计算两个用户评分向量可表示为如下:

其中，表示用户u和用户v都评分过的项目集合，而用户u和用户v评分的均值则分别用与表示。

通过上述的相似度计算公式，计算系统中所有用户的相似，可得到如下相似度矩阵

S:

通常情况下S为对称矩阵，即元素之间存在。矩阵S元素表示用户u与用户v之间的相似度，即。利用相似度矩阵，便可以为目标用户选取一定数目的近邻。近邻选取的方式可以分为:固定值法和阈值法两种。固定值法，又称固定近邻个数法，其是为目标用户选择与其相似度最高的k个用户作为最近邻居;而阈值法，又称相似度阀门法，其是选择那些与目标用户的相似度大于某个值的用户作为最近邻居。这两方法各有优缺点，固定值法能找到足够的用户作为近邻，但可能因为近邻与目标用户的相似度不高，而影响最终推荐结果;相反阈值法由于可以控制近邻与目标用户之间的相似度，固选取的近邻与目标用户之间的相似度一般较高，但可能因为阈值设定的过高，而不能得到足够多的近邻，而影响推荐的质量。

3.3为用户产生推荐

当目标用户的最近邻选取完后，就可以根据基于协同过滤算法的核心思想，具有相似兴趣偏好的用户，其对信息的需求也存在一定的相似性，来对目标用户对未知项目的评分进行预测，并最终将推荐结果反馈用户。[8]加权平均值公式如下:

其中，分别为用户u与用户v项目的评分的平均值。

## 3.2 技术路线

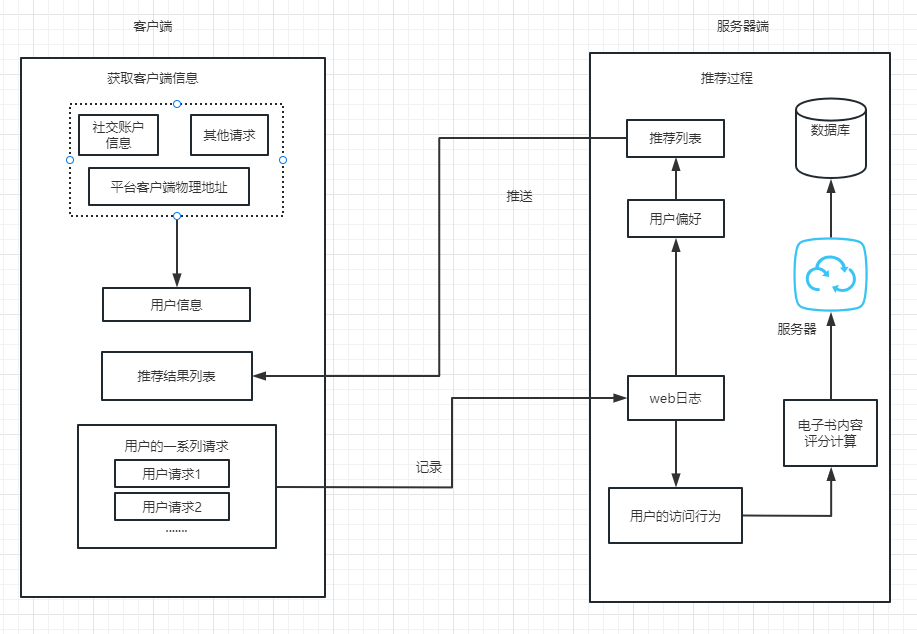


图3-1 技术路线图

系统基于Android Studio平台使用Java语言进行开发，目前设计思路为采取前后端分离式系统设计。

安卓App：只承担前端收集用户的信息和展示内容的任务。用户登录时输入邮箱和密码，传入到登录用的 Login(String email,String password)方法，建立User实体类的实现，进行登录操作。 APP获取用户的大量电子书数据，包括书名、作者、出版社、分类、评分、评论以及购买/借阅记录等信息，并对数据进行筛选、清理、标准化和统计，以满足后续的建模需求，并通过OKHttp3创建OkHttpClient对象使用异步POST请求封装用户相关信息，并向后端服务器发送请求。接收到服务器的请求回复后，解析数据，通过ListView和ImageView等组件展现给用户。

后端服务器：采用Python Flask技术搭建一个轻量级服务器，承担解析用户的请求，并训练和调用深度学习模型以及数据库的任务。后端服务器在接收到请求后进行解析，将原始数据转换成计算机可识别的特征，例如文本特征（TF-IDF、Word2Vec）、图像特征（SIFT、SURF）等，在这个过程中，需要挖掘与预测目标相关的重要特征，因此采用特征选择技术。之后通过调用Pytorch库的相关函数并结合相关的推荐算法，对上述特征进行训练和优化，并根据不同指标（如准确率、召回率、F1-Score、AUC等）对模型进行评估和调优。根据用户的购买历史、评价和兴趣偏好，通过上述模型预测用户可能感兴趣的电子书，并将推荐结果通过异步POST请求以列表、瀑布流、卡片等形式呈现给用户。不断收集用户反馈和数据，根据用户行为和模型表现调整模型参数，优化推荐策略，从而实现不断改进和提高推荐系统的性能。

数据库设计：本系统采用MySQL数据库存储用户、标签、用户标控、图书、图书标签、图书评分、用户图书标签等实体和实体之间的关系。其中用户表存储用户的姓名、邮箱、密码等信息；标签表存储标签名称、标记次数等信息；用户标签表存储用户感兴趣的标签、标签标记的次数等信息；图书表存储书号、书名、内容简介、出版社等信息；图书标签表存储图书与标签之间的关系以及被标记的次数等信息；、图书评分表存储图书的评分记录等信息；用户图书标签存储用户对喜欢的图书的标签、评分等信息。

# 4 研究难点及目前存在的问题

## 4.1 研究难点

冷启动问题, 稀疏性问题, 兴趣转移问题,扩展性问题

## 4.2 目前存在的问题

1.冷启动问题包括新项目问题与新用户问题。新项目问题是指对于新加入系统的项目，还没有任何用户对它评分，如果单纯的依靠用户历史的协同过滤，则新物品永远不会得到推荐，显然这是我们不愿意看到的，因此我们需要一定的策略进行新项目的推荐。新用户问题是指协同过滤在学习一个用户的兴趣时是学习该用户的行为记录，但对于没有任何记录的新用户而言，就无法学习用户的兴趣，并与其它用户进行比较，从而不能做出正确的推荐。

2.稀疏性问题：由于每个用户的精力是有限的，不可能对所有物品评分，大多数用户都只会对少部分的物品进行评价，对于任何一个推荐系统而言，已评分数据都要远小于未评分数据，即数据十分稀疏。在刚建立系统时，这个问题尤其严重。

## 4.3 拟采取的措施

1.解决新项目问题的一般方法是利用物品的内容信息，通过基于内容的预测评分，来弥补新物品与已有物品之间的鸿沟。内容信息是物品本身所固有的，所以不会受到没有用户评分的牵制。[10]

对于新用户问题，有些推荐系统在新用户注册的时候会随机提供给用户一个项目集合，并要求用户对集合中的项目进行评价，例如著名的豆瓣网就要求用户注册时评价一定数量的内容。这种方法在一定程度上解决了新用户问题，但它需要用户的配合，有时用户可能对网站所提供的项目内容并不了解，无法做出准确的评价。

2.在统计学中，采用最大似然估计的方法，可以对不完整的样本数据进行学习，得到一个能够很好拟合观察样本数据的模型，从而可以适应数据的稀疏性。

# 5 研究预期成果

1. 实现一个高效、准确的电子书推荐系统，提供用户个性化的阅读体验，提升用户的阅读满意度和良好评价率。

2. 通过数据挖掘、机器学习等技术，为用户智能推荐具有吸引力和实用价值的电子书，改善用户的阅读体验和阅读效率。

3. 深入挖掘用户行为数据，分析用户的阅读偏好和需求，优化推荐算法，提高推荐的精准度和个性化水平。

4. 结合社交网络、内容标签等多种信息源，构建多元化的推荐体系，扩大推荐的范畴和覆盖面，增加用户的选择和满意度。

# 6 进度安排

第1周，联系导师，深入沟通。项目准备及调研，学习与课题相关的基础知识，掌握课题相关知识。  
第2周，下达任务书，明确任务需求，毕业设计正式开始。  
第3-4周，需求分析和概要设计。仔细阅读毕业设计任务书，完成需求分析和概要设计，写出需求分析和概要设计报告。  
第5-8周，详细设计和部分编码。在需求分析和概要设计的基础上，完成系统的详细设计，写出详细设计报告，同时完成部分编码工作  
第9-13周，中期检查；编码与实现。进行软件编码和系统的具体实现， 编写出源代码。其中第9周，学院对毕业设计进行全面中期检查，检查采取现场检查的方式。需填写《本科毕业设计（论文）中期检查调查表》。  
第14-15周，系统测试和试运行。经测试能正常可靠地运行。整理资料，写出论文，毕业生论文打印提交，提交毕业论文，准备预答辩和查重。  
第16周，预答辩，论文查重及论文质量抽查。导师安排一次预答辩，并对论文格式、内容等把关；学院进行论文查重，对论文质量进行抽查。  
第17周，毕业设计（论文）答辩。指导教师填写评语、论文评审；毕业答辩，整理毕业设计资料，交导师存档。答辩通过，争取取得优良成绩。

**文献总结**

[1] 论文介绍了主流的推荐算法基于内容的推荐算法、协同过滤算法等。论文在综合了上述主流推荐算法的优缺点的基础上，提出一种基于内容和协同过滤的混合推荐算法（CB-TUB)，以此来提升算法性能。论文的具体工作包括:第一，针对项目“冷启动”问题，用特征“标签”对项目内容进行描述，进行基于内容的推荐。第二，在利用基于用户的协同过滤推荐算法进行推荐时，通过用户对项目的隐式评分来构建用户-评分矩阵，以此缓解数据稀疏性问题。在相似度计算时，一方面引入用户的评分时间权重，体现用户兴趣变化;另一方面，引入杰卡德统一算子距离，衡量用户间共同评分项目所占的比重。最终，论文创造性地提出了“用户感兴趣指数”这一指标以便两种算法能在同一维度上进行混合，并给出了计算方法。在算法混合时，将推荐项目的加权后用户感兴趣指数从高到低排序，为用户产生Top-N推荐。

[2] 该文献是 Paul Resnick 和 Hal R. Varian 在 1997 年发表于《ACM通讯》杂志上的一篇文章，题目为“推荐系统”（Recommender Systems）。文章介绍了基于协同过滤和基于内容过滤两种推荐算法的不同之处、优缺点和应用场景，为推荐系统领域奠定了基础。

该文献中，作者首先介绍了什么是推荐系统、推荐系统的应用场景和推荐系统的发展历程。随后，作者分别介绍了基于协同过滤和基于内容过滤两种推荐算法，并对其进行了比较和分析。基于协同过滤的推荐算法适用于有足够用户行为数据的情况下，能够利用用户的历史行为找到与之相似的其他用户并据此进行推荐。而基于内容过滤的推荐算法则是根据物品的属性和特征来进行推荐，相对于基于协同过滤，其具有数据稳定性强、算法简单等优点，但需要预先对物品的属性和特征进行建模和提取。此外，作者还提到了混合推荐的概念，即将不同的推荐算法组合起来使用，以取长补短。

[3] 这篇文献是由 Nishanth Lathia、Licia Capra 和 Shane Hailes 在 2008 年发表在 ACM 推荐系统会议上的论文，题目为“KNN CF: A Temporal Social Network”。该论文提出了一种基于 KNN 算法和时间信息的协同过滤推荐算法，用于解决传统的基于协同过滤算法在处理动态社交网络数据时的困难。

在文献中，作者说明了传统的基于协同过滤算法在处理动态社交网络数据时的局限性，包括利用用户的静态行为数据进行计算容易产生冷启动问题、对于过于活跃的用户会出现过拟合现象等。为了解决这些问题，作者提出了基于 KNN 算法和时间信息的协同过滤推荐算法。该算法通过对用户行为数据进行时间间隔分析，将相邻不同时期的行为数据分开进行处理，避免不同时期行为数据之间的影响，同时运用 KNN 算法对相似用户进行预测，解决了过度活跃和过拟合问题。该文献中提出的算法不仅在处理动态社交网络数据的情况下具有较好的效果，而且在算法实现以及扩展性方面也具有优越性。

[4] 该文献由 Ignacio Siles、Johan Espinoza-Rojas、Adrian Naranjo等一组作者在2019年发表于IEEE Transactions on Emerging Topics in Computing, 文章题目为“The Mutual Domestication of Users and Algorithmic Recommendations on Netflix”。

该文献探讨了Netflix推荐系统如何通过算法和用户之间的相互适应来提高推荐的效果。文献作者分析了Netflix的大量用户数据，并通过实验来探究用户与Netflix推荐算法之间的相互作用。实验表明，Netflix的推荐算法与用户之间的关系不是单向的，而是相互作用的。Netflix的推荐算法会影响用户选择的影片和观看习惯，同时，用户的选择和观看习惯也会影响并调整推荐算法。因此，两者之间的相互适应是导致Netflix推荐算法成功的关键。该文献给出了一种全新的观点，即对于个性化推荐，需要把推荐算法看作是一种“陪伴者”，而不是一种“服务提供者”。它们需要与用户建立一种“互动策略”，以相互适应和提高推荐的效果。

[5] 该文献是由陈功平和王红在2016年发表于《山东农业大学学报(自然科学版)》上的一篇论文，题目为“改进Pearson相关系数的个性化推荐算法”。该论文提出了一种基于改进的Pearson相关系数的个性化推荐算法，用于提高用户对推荐结果的满意度。

在文献中，作者指出传统的基于协同过滤的推荐算法存在冷启动问题、数据稀疏问题等缺点，并提出了基于改进的Pearson相关系数的个性化推荐算法。在算法实现时，首先利用改进的Pearson相关系数计算出用户之间的相似度，并预测出用户对于未评分物品的兴趣度，在此基础之上进行物品推荐。与传统的Pearson相关系数相比，改进后的Pearson相关系数加入了用户评分数量的因素，提高了相似度计算的精度和推荐结果的准确性。

该文献中提出的基于改进的Pearson相关系数的个性化推荐算法，相较于传统的协同过滤算法有更好的适用性和有效性。该算法考虑了用户评分数量与用户相似度之间的关系，克服了数据稀疏和冷启动问题，提高了推荐结果的准确性，并可以根据用户的历史数据动态调整推荐策略，实现个性化推荐。

[6] 在文献中，作者指出传统的推荐算法对于用户兴趣的建模过于简单，无法解决复杂数据场景下的推荐问题，因此提出了一种基于图网络结构的推荐方法。该方法将用户、物品以及评分之间的关系构建成一个图网络，并利用图网络中的节点和边来表示用户和物品之间的关系。同时，该方法结合了图网络分析和深度学习方法来实现对用户兴趣的建模，并利用推荐技术对用户进行推荐。

该文献中提出的基于图网络结构的推荐方法，通过模型的建立和训练，能够动态更新用户的兴趣模型，解决了传统推荐算法对于用户兴趣建模过于简单的问题，并且具有较高的推荐准确性和推荐效果。

[7] 在文献中，作者指出传统的协同过滤算法存在数据稀疏和冷启动问题，同时推荐的物品也会出现类似的问题，因此提出了一种基于协同过滤和关联规则的混合推荐算法。该算法通过结合协同过滤和关联规则两种算法，综合利用用户行为数据和物品属性信息进行推荐，提高了推荐准确性和覆盖率。

在算法实现时，该文献中提出了一种基于 Apriori 算法的关联规则挖掘方法，用于发掘用户的关联购买行为，同时利用物品属性信息来对推荐结果进行过滤，从而提高了推荐的准确性。

该文献中提出的基于协同过滤和关联规则的混合推荐算法，通过综合利用不同的推荐算法的优势，可以克服传统推荐算法的诸多缺点，提高了推荐准确性和覆盖率。同时，该算法还提出了一种基于关联规则的物品过滤方法，能够更好地满足用户的个性化需求，提高了推荐的质量和用户满意度。

[8] 在文献中，作者指出传统的推荐算法无法有效挖掘用户的情感信息和文本评价信息，导致推荐的物品和服务与用户的个性化需求不匹配，因此提出了一种基于文本情感分析和矩阵分解的混合推荐算法。该算法结合了文本情感分析和矩阵分解技术，利用用户的情感和文本评价信息，对用户的兴趣进行建模，并进行个性化推荐。

在算法实现时，该文献中提出了一种基于情感词典和情感分析模型的情感分析方法，用于提取用户的情感信息，同时结合矩阵分解技术建立用户兴趣模型，并通过推荐技术实现个性化推荐。

该文献中提出的基于文本情感分析和矩阵分解的混合推荐算法，能够更好地利用用户的情感和文本评价信息，提高了推荐的准确性和个性化，满足了用户对于推荐的个性化需求。

[9] 该论文主要研究了基于协同过滤的个性化推荐算法，包括用户协同过滤算法、物品协同过滤算法和基于矩阵分解的协同过滤算法等。

在用户协同过滤算法中，作者提出了一种基于领域相似度和用户相似度的加权用户协同过滤算法，通过对用户评分进行加权计算，提高了推荐的准确性和效率。在物品协同过滤算法中，作者提出了一种基于物品相似度和用户历史行为的加权物品协同过滤算法，通过对用户历史行为进行加权计算，提高了推荐的精度和召回率。

另外，作者还提出了一种基于矩阵分解的协同过滤算法，该算法通过将用户物品评分矩阵分解为两个低维矩阵，以降低数据维度和稀疏性，提高推荐准确性和效率。同时，该算法还能够适应动态变化的用户兴趣和物品特征，提高了算法的实用性和鲁棒性。

最后，作者通过实验对比了不同算法的推荐效果和性能，验证了所提出算法的有效性和优越性。

[10] 浏览新闻已经是大多数人的生活习惯,如何从海量的新闻中获取感兴趣的新闻，成为很多人关心的问题，可以节省大量时间，新闻推荐系统可以完美解决这个问题。文章从基于协同过滤的推荐算法入手，通过计算余弦相似度来获取用户的相关特征。并给出了基于用户和基于物品的两种协同过滤算法的对比，分析了两种算法在不同场景下的优劣。并在最后对以上两种算法进行了实验验证。该算法首先将用户和新闻分别抽象成矩阵，然后使用协同过滤算法对矩阵进行分解，得到用户和新闻的特征向量。接着，通过计算用户和新闻特征向量的相似度，将相似度高的用户和新闻匹配推荐给用户。文章还结合实例展示了算法的具体实现和推荐效果。比较实验显示，该算法在推荐准确率、召回率等指标上都有着不错的表现，可以提高用户对新闻推荐的满意度。最后实验表明基于用户的协同过滤算法在准确度上略微高于基于物品的协同过滤算法。

[11] 该文从信息过滤技术引出个性化的推荐系统，利用用户的信息、web内容以及用户行为，推测当前用户可能感兴趣的内容，然后进行推荐。并分析了基于协同过滤的算法

存在的一些问题，以及解决手段。将内容过滤与协同过滤相结合解决稀疏性和冷启动的问题，并引入了贝叶斯网络模型来实现内容过滤，解决了冷启动问题。又采用NMF的协同过滤以适应稀疏的数据，然后将两者结合起来。采用线性逐渐遗忘的方法来调整用户评分的权重，以解决大多数传统的协同过滤研究所忽略的用户兴趣转移问题。

# 参考文献

1. 胡晓雪.基于内容和协同过滤的混合推荐算法研究及应用[D].重庆大学,2020.DOI:10.27670/d.cnki.gcqdu.2020.002555.
2. Resnick P, Varian H R. Recommender systems[J]. Communications of the ACM. 1997,40(3):56-58.
3. Lathia N,Hailes S, Capra L. KNN CF: a temporal social network[C].ACM Conference onRecommender Systems.ACM, 2008:227-234.
4. Siles Ignacio, Espinoza-Rojas Johan, Naranjo Adrian, et al. The Mutual Domestication of Users andAlgorithmic Recommendations on Netflix. 2019,12(4):499-518.
5. 陈功平,王红.改进Pearson相关系数的个性化推荐算法[J].山东农业大学学报(自然科学版).2016,47(06):940-944.
6. 王睿.基于图网络结构的推荐方法研究[D].哈尔滨理工大学,2019.
7. 刘晓琳.基于协同过滤及关联规则的混合推荐算法研究[D].西安电子科技大学,2014.
8. 吴晓亮.基于文本情感分析与矩阵分解的混合推荐算法研究[D].南昌大学,2019.
9. 叶卫根. 基于协同过滤的个性化推荐算法研究[D].江南大学,2016.
10. 李梅.基于协同过滤的新闻推荐算法研究[J].电脑知识与技术,2022,18(34):51-53.DOI:10.14004/j.cnki.ckt.2022.2228.
11. 段玮. 基于协同过滤的个性化推荐算法研究[D].华中科技大学,2009.