分类号 TP391 学校代码 10590
UDC 004 密 级 公开

深圳大学硕士学位论文

基于深度神经网络的视频个性化推荐系统研究

学位申请人姓名 高睿
专 业 名 称 信息与通信工程
学院（系、所） 信息工程学院
指导教师姓名 李霞 教授 陈亮 博士

深圳大学学位论文原创性声明和使用授权说明

原创性声明

本人郑重声明：所呈交的学位论文 基于深度神经网络的视频个性化推荐系统研究 是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含任何其他个人或集体已经发表或撰写过的作品或成果。对本文的研究做出重要贡献的个人和集体，均已在文中以明确方式标明。本声明的法律结果由本人承担。

论文作者签名：高肯　　日期：2017年5月13日

学位论文版权使用授权书

本学位论文作者完全了解深圳大学关于收集、保存、使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权单位属深圳大学。学校有权保留学位论文并向国家主管部门或其他机构送交论文的电子版和纸质版，允许论文被查阅和借阅。本人授权深圳大学可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（涉密学位论文在解密后适用本授权书）

论文作者签名：高肯　　导师签名：李晨

日期：2017年5月13日　　日期：2017年5月13日

摘　　要

随着信息化时代的到来，视频数据量与日俱增。一方面视频供应商希望能获取用户的兴趣偏好，投其所好为用户推送他感兴趣的视频，提高用户体验，增强用户对公司的粘合度；另一方面，网络视频用户希望能从海量视频内容中，快速地找到自己感兴趣的的视频，减少不必要的时间开销。正是这些迫切的需求，使得视频个性化推荐成为了一个热门的研究课题。在这个背景下，本文对视频的个性化推荐算法展开了研究，并且利用深度学习在自然语言处理上的相关技术，提出了两种基于深度神经网络的视频个性化推荐算法：

1）基于深度语义模型的视频个性化推荐算法。该算法结合了深度学习和基于内容的推荐算法，通过构建一个深度网络模型，对视频和用户的本文信息进行特征提取，并在语义空间层面完成对视频和用户的分布式特征表示，深度挖掘用户和视频之间的潜在联系，并进行视频推荐。

2）基于概率语言模型的视频个性化推荐算法。该算法首先将用户历史观看所得到视频序列，类比成自然语言序列，将序列中的每个视频项看做是语句中的独立的单词，然后结合深度学习在自然语言处理上的相关技术，根据视频与视频之间以及用户与视频之间的相关性大小，构建并训练神经网络模型，得到用户与视频的分布式特征表示。并在此基础上，利用视频与用户向量之间的相似性进行视频推荐。

在文中，首先通过理论推导分别对所提出两种推荐算法的原理进行了详细阐述，在理论层面分析它们的可行性。然后根据算法的相关原理，通过对采集到的视频中的用户历史点击行为数据以及视频文本描述数据进行了实验设计。并将所得的结果分别与 TF-IDF 、UserCF、ItemCF 等算法的结果进行了对比分析。实验结果表明，本文所提出的两种推荐算法的结果都要优于对比算法，都能够较好地完成视频推荐任务。

关键词：信息过载；个性化推荐；深度神经网络；自然语言处理；协同过滤

Abstract

With the coming of the information age, the amount of video data is increasing day by day. On the one hand, video providers hope to obtain user preferences, match up to the user to push his interest in video, improve the user experience, enhance the user stickiness of the company; on the other hand, online video users hope from the massive video content, quickly find themselves interested in the video, reduce unnecessary time overhead. It is these urgent needs that make personalized video recommendation become a hot research topic. In this background, this paper studies the personalized recommendation algorithm of video, and proposes two kinds of personalized recommendation algorithms based on the deep learning technology in Natural Language Processing.

1) Video personalized recommendation algorithm based on depth semantic model. The algorithm combines depth learning and content-based recommendation algorithm, and constructs a depth network model to extract the feature of video and user's information. At the semantic level, the distributed representation of the video and the user is realized, and the potential connection between the user and the video is deeply excavated.
2) Video personalized recommendation algorithm based on probability of language model. The algorithm first will get the video sequence, the user history to watch the analogy to natural language sequence, the sequence of each video as a statement of the separate words, Then combined with the depth study on natural language processing technology, according to the correlation between different video size, building and training the neural network model, distributed characteristics of the video said. And on this basis, using video vector similarity between collaborative filtering recommendation.

Respectively in this paper, first of all, through the theoretical derivation of the proposed two kinds of recommendation algorithm, this paper expounds in detail the principle of the

theoretical analysis of the feasibility of them. Then according to the related principle of the algorithm, based on the collected the user history click behavior of Tencent video data and video text description data has carried on the experimental design. And converting the proceeds respectively with the results of the UserCF, ItemCF, TF-IDF algorithm compares the results of analysis, etc. The experimental results show that the presented two kinds of recommendation algorithm is superior to contrast the result of the algorithm, they can be very good video recommendation task.

Key word： Information overload；Personalized recommendation；Deep neural network；Natural language processing； collaborative filtering

目　录

第1章 绪 论

1.1 研究背景及意义

近年来，在互联网技术飞速发展的大背景下，信息数据呈指数形式式增长，每天都有大量的数据在不断地更新和产生，例如，新浪微博上的各种动态更新，腾讯视频库里数以亿计的视频，淘宝、Amazon 上琳琅满目的商品，Delicious 上超过 10 亿的网页收藏等[1]。人们已经开始从一个信息缺乏的时代进入到了一个信息爆炸的时代。信息的冗余，无论是对信息的消费者，还是对其生产者而言，都是一个新的挑战。对前者而言，如何在海量的信息资源中通过高效地筛选获取对自己有用的信息，成为了信息消费者将有面临的一道难题；而对于后者来说，如何快速精准地找到不同用户群体的兴趣喜好，投其所好提供用户所感兴趣的内容，增强用户体验，加强用户对自己产品的粘合度，这也成为他们迫切希望解决的问题。

搜索引擎[2]是人们为了解决信息过载问题而提出来的一种解决方案。当人们对自身所需要的目标信息比较明确时，人们可以通过在搜索引擎中输入一组关键词的来进行信息检索，然后服务器根据相关算法计算出各个信息项与用户所需信息的相似度，最后按相关性大小进行排名，并将排名结果返回给用户。尽管搜索引擎在一定程度上使得用户获取信息的效率得到了较大提升，但是它有着很大的局限性。这是因为搜索引擎的正常运作的前提是要求用户对自己所需信息比较明确，它需要用户能够通过几个关键来描述自己所需的信息。然而在日常生活中，很多情况下人们对自己所需的信息并不是很了解，很难搞将自己的信息需求通过关键词来进行描述，这时搜索引擎就无法帮助人们获取准确信息了。

而推荐系统[3]的出现很好地弥补搜索引擎的不足，它无需用户主动对自己的信息需求进行描述，而是根据用户的历史行为数据分析用户的兴趣偏好，找到他所感兴趣的对象集合并将结果反馈给用户。 与传统的搜索引擎不同的是，个性化推荐系统主要是通过利用用户的历史行为信息以及用户的历史评分信息等数据来分析和查找用户的兴趣偏好，因而可以启发式地帮助用户找到其所求，达到具有针对性地的推荐的目的。此外，

个性化推荐结果是随着用户历史信息的变化而时刻改变的，也即是说当用户的行为数据出现改变的时候，个性化推荐的结果也会相应地发生变化。

个性化推荐技术的理论思想是于 20 世纪末首度产生并逐渐流行起来的，主要是得益于电子商务的兴起和发展。1997 年，Resnick 等[5]人对推荐系统给出简单的定义：“它是利用电子商务网站向客户提供商品信息和建议，帮助用户决定应该购买什么产品，模拟销售人员帮助客户完成购买过程”。

近些年来，个性化推荐系统成为了互联网各个行业中所研究的热门课题，并且已成功地应用到了诸多领域。特别地是在视频领域，随着视频多媒体技术的快速发展，每天都有海量视频信息上传到视频网站中，视频信息量的激增必然要面临非常严重的信息过载问题。使得个性化推荐服务在视频播放的意义十分重大，其作用显得尤为重要，且其地位不可替代。

目前，视频网站大致可以分为两类：一类是以互联网用户自主上传原创内容为主的视频网站，如 YouTube、优酷、Bilibili 等；另一类则是以提供某些影视作品为主的视频网站，如腾讯视频、爱奇艺等。对比而言，第一类网站的视频数目相对较多，其展现的内容也更为丰富多样，但原创视频的生命周期会相对较短，同时，视频的质量水平不一，且以短视频占大多数，数据结果参差不齐。而第二类视频网站由于其提供的更偏向于专业的影视内容，因而具有较好的结构化数据，但内容则相对比较单一。无论是哪类一类视频网站随着视频资源不断地增加，必然将会面临信息冗余的问题。一方面，视频用户期望能快速地从海量视频资源中定位到个人喜爱的内容，提高获取视频的效率，另一方面，企业也希望能准确定位到对特定视频内容兴趣度高的目标用户，做出人性化的推荐，以提升其对企业视频网站的忠实度和依赖度，减少客户流失。为此，现在在优化视频系统领域的研究中，寻找更合适有效的方法以提高为特定用户推荐符合其偏好视频内容的精确度，已经成为了一个非常重要的研究内容。

视频个性化推荐系统的最大的特色在于：第一，推荐系统所提出的个性化推荐决策是一种主动行为，它是从用户的历史观看行为出发，结合相应的推荐算法，最后自主地完成推送；第二，个性化推荐系统做出的推荐决策是不断实时更新的，系统所生产的特定用户的推荐列表也是随着数据库中用户偏好数据和其对应的视频评分数据的更新而发生变化的。因此，视频个性化推荐系统能够显著地提升视频推送的高效性，提高企业服务质量，带来更优质的用户体验，实现用户和企业的双赢模式。

在实际应用中，对于商家、用户、学者以及整个信息社会而言，视频的个性化推荐

系统的研究拥有很高的经济效益价值、学术科研价值和社会效益价值。因此，个性化视频推荐系统的研究、应用和发展对经济社会发展的很多方面都有着重要的意义，也必将对人们信息需求方面的生活方式产生深刻的影响。

1.2 国内外研究现状

国外对个性化推荐系统的研究可以追溯到上个世纪 90 年代，自从第一篇关于协同过滤算法的文章问世以来[6]，个性化推荐系统便开始成为了一个非常热门的研究课题，并且在过去数十年里，工业界和学术界取得了大量的研究成果。

最早关于个性化推荐的理论研究出现在 1992 年，当时 Goldberg 等[3]人提出的一种基于协同过滤的推荐算法，并且他们将该算法应用到了新闻的个性化推荐中。该算法首先通过分析用户的历史点击行为记录，以此计算出用户之间兴趣的相似度，最后利用不同用户之间的相似度进行新闻推荐。在同年，美国的 Xerox 公司利用协同过滤算法提出并设计了 Tapestry 系统[4]，并将该系统应用在了邮件推送领域，通过用户反馈的历史信息，分析用户的喜好构建用户模型，利用协同过滤算法设计了邮件的过滤系统，这是协同过滤的最早实践。紧接着在两年后，美国明尼苏达大学的相关研究人员提出了一款名叫 GroupLens 的新闻的个性化推荐系统[7]，该系统首先要用户根据自己的喜好对浏览过的新闻进行打分，然后系统通过这些评分记录分析用户的兴趣偏好，计算用户彼此之间的兴趣相似度，进而给用户进行推荐。在此基础之上，他们后来还提出了性质相近的 MovieLens 电影推荐系统，同样是利用协同过滤算法，分析用户的喜好，找寻相似的用户，以此作为依据进电影推荐。

上述三个事件被许多学者看做是推荐系统研究开始的标志性事件。在后来时间里，研究人员又陆陆续续地发表很多优秀的研究成果。例如，Lemire 和 Maclachlan 等[8]人提出了 Slope One 算法针对协同过滤推荐算法中评分矩阵稀疏性的特点，对其进行适当的填充，并取得了较好的推荐效果。Chang 等[9]人通过在传统的协同过滤算法中应用神经网络算法，也达到了提高推荐质量的目的。MIT Media-Lab 提出了一种基于社会网络模型的推荐算法，在此基础上他们开发出了一款音乐推荐系统 Ringo[10]，该系统通过对用户的社会信息进行过滤，构建用户的社会化网络，然后再进行音乐推荐，并且取得了不错的推荐效果。Jonghun 等[11]人提出一种基于文本挖掘的视频推荐算法，该算法通过利用文本挖掘领域的相关技术，首先将视频表示成关键词向量，然后将用户所观看的视频向量求得算术平均生成用户向量，最后通过计算用户向量和视频向量之间的相似度生成

推荐结果反馈给用户。除此之外，在 2007 年的时候，还成立了专门的推荐系统会议（Re-cSys），这是推荐技术研究和应用最高级别年度会议。

国外的个性化推荐研究不仅在理论研究方面取得了长足的进步，也在实际应用中取得了丰硕的成果。当前已有多家企业成功的将个性化推荐系统应用到自己的业务中，最为著名的有 Amazon 的书本推荐系统、eBay 的商品推荐系统。同时，这些成功的应用范例也反过来进一步的推动个性化推荐技术的革新发展。

国内对个性化推荐系统的研究要晚于国外，当使用中国知网学术趋势搜索以“个性化推荐”为关键字进行检索时，便可以得到如图 1-1 所示的个性化推荐学术关注度的趋势图。从图中我们可以看到，国内对推荐系统的研究的真正起步时间是 2004 年，从那以后个性化推荐在国内的关注度才开始显著提高，然后到 2010 年之后，其关注度开始飞速增长，并呈指数增长趋势。不难理解，从 2004 至今的数十年正是互联网行业在国内高速发展的数十年，信息时代的发展不可避免地会带来严重的信息过载问题，因此推荐系统的研究具备了很强的现实意义，使得越来越多的国内学者和企业家逐渐意识到了个性化推荐的重要性。

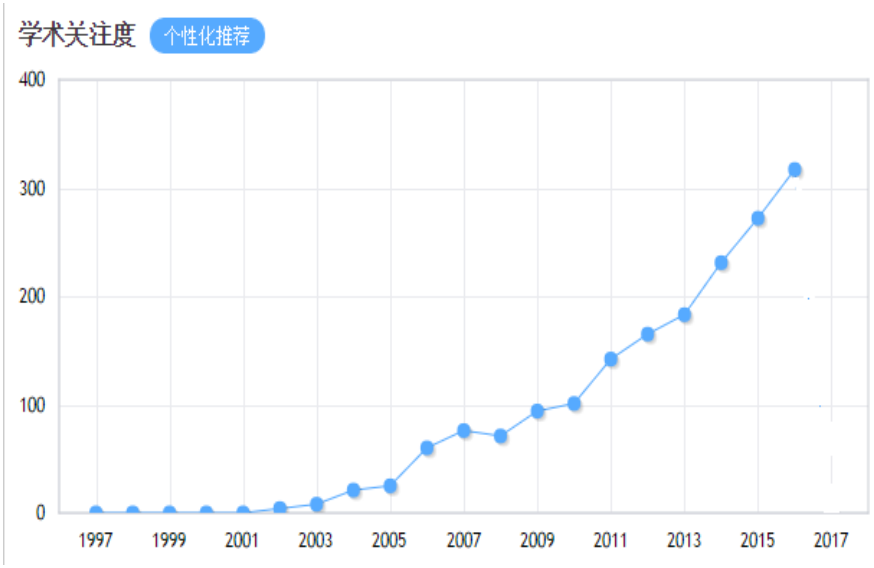

图 1-1 个性化推荐学术关注度趋势图

目前，国内已有大部分的电子商务平台都在逐步构建自己的个性化推荐系统，如淘宝、京东等，它们都拥有自己的个性化推荐系统以满足不同用户的特定需求。此外，有很多功能性网站，如爱奇艺、腾讯视频等推出了视频的个性化推荐服务，帮助用户高效地寻找自己感兴趣的视频节目；新浪微博等推出了基于人与人的推荐功能，把握个性化推荐服务的契机，增加用户的忠实度；今日头条等提出了基于个性化推荐的连接人与信

息的新型服务模式，成为了互联网行业近年来的一匹黑马，正在迅速地发展壮大。

除此之外，在学术上，国内的对个性化推荐系统的研究也获得了不错的成绩。徐翔等[12]人针对协同过滤推荐系统，提出了一种相似度度量的优化算法，提高了推荐的精度。尹路通等[13]人提出了一种基于利用隐语义模型的视频推荐策略，挖掘对用户与网络视频之间的隐含特征关联，该算法，在一定程度上降低了人工成本和时间开销，并且提高了用户兴趣建模的准确度。杨兴耀等[14]人针对协同过滤中的数据稀疏性问题，通过对推荐物品进行人工标注，丰富物品的特征属性，深层次地发掘用户-属性、属性-项目之间的关联关系。

1.3 论文主要研究内容及章节安排

1.3.1 主要研究内容

深度学习是近年来非常热门的研究方法，它在图像处理、语音识别、自然语言处理等领域都取得了巨大的成就，然而目前基于深度学习的推荐算法的研究与应用仍在探索阶段。深度学习具备强大的特征表征能力，它能够对低层特征进行提取，生成抽象的高层特征，以发现输入数据分布式特征信息。因此，本文利用深度学习的这一特性，对基于深度学习的视频个性化推荐算法展开。本文主要的工作内容包括：

1）深入地研究和学习了当前主流推荐算法的相关原理，并且总结与分析了各自算法所存在的优缺点。

2）提出了一种基于深度语义模型的视频个性化推荐算法。该算法针对协同过滤数据稀疏性和冷启动问题结合了深度学习与基于内容的推荐算法对视频的个性化推荐展开研究。在该算法中通过构建的深度神经网络模型，对视频和用户的文本内容进行特征提取，获取用户与视频的语义特征，深度挖掘用户与视频的内在相关性，然后根据基于内容推荐的相关策略进行视频推荐。

3）提出了一种基于概率语言模型的视频个性化推荐算法。该算法受到概率语言模型的启发，将用户所观看的视频序列类比成自然语言，并通过分析视频之间以及用户与视频之间的关联信息建立数学模型。然后搭建神经网络训练加入用户信息的视频序列，提取用户和视频的特征，得到用户与视频的分布式特征表示。然后利用用户与视频向量之间的相似性生成推荐列表，并将结果反馈给用户。

4）针对文中所提出的视频推荐算法，设计了实验，对算法进行实践。并将实验结果与常用的推荐策略进行了对比分析，总结了算法的优缺点，并最后提出改进的方向。

1.3.2 章节安排

基于以上的工作内容，本文将各章节安排如下：

第 1 章为论文绪论部分。首先是介绍个性化推荐系统的研究背景，并且探讨其的研究意义。然后阐述个性化推荐系统在国内外的研究现状，最后简要概述本文主要的研究内容。

第 2 章为论文的理论知识部分。主要介绍个性化推荐系统的相关理论知识，然后根据推荐算法的不同将推荐系统进行了分类，并对不同类别的推荐系统进行了介绍，最后分析了各自优缺点。

第 3 章主要介绍基于深度语义模型的视频个性化推荐算法。首先对该系统的框架进行描述，然后通过理论推导对其原理进行详细阐述，然后根据算法流程设计实验，最后对实验结果进行展示和分析。

第 4 章主要介绍基于概率语言模型的视频个性化推荐算法。首先根据概率语言模型提出了一种用户以及视频向量化的算法，并阐述该算法的基本原理，对其进行理论推导。然后基于该算法所生成的视频向量以及用户向量，计算视频与用户之间的相似度，以此作为依据进行视频推荐。最后根据算法流程设计实验，并对结果进行对比分析。

第 5 章为本文的工作总结部分。简要探究了本文研究内容的局限性，并提出了对未来研究工作的改进方法。

第 2 章　个性化推荐系统及相关技术

在科技不断高速发展的今日，视频业务已然成为了与人们生活密不可分的服务之一，视频业务后续的持续发展，离不开有效的推荐技术，而高效率高节奏的生活，也使得人们迫切的需要推荐技术能帮助他们从海量的视频数据库中精确高效地筛选出自己感兴趣的资源，因而，研究推荐系统及其相关技术，具有重大的现实意义及研究价值。

2.1 推荐系统概述

推荐系统主要是用于解决日趋严重的信息过载问题，它通过利用用户的历史行为数据，分析用户的兴趣爱好并构建相应的用户模型，从待推荐的项目中选择与其兴趣偏好相符的项目进行推荐。它通过分析海量日志信息，构建用户与项目之间、或项目与项目之间、异或用户与用户之间的二元关系，以此分别建立用户与视频模型，利用不同内容之间的相似性实现个性化推荐。

Goldberg 等[3]人曾给出了个性化推荐系统的通用模型，如图 2-1 所示。从图中可以看到，推荐系统的工作流程大致可以概括为：1)，获取用户的兴趣偏好，其获取方式主要有两种：第一种是由用户自己主动提供兴趣偏好，这是一种显示的获取过程；另外一种则是由推荐系统通过分析用户历史行为数据，通过相关算法挖掘出用户的兴趣偏好，这种获取过程是隐式的；2)，推荐系统根据所获得的用户的兴趣偏好对用户进行模型构建；3)，根据建好的用户模型结合相应的推荐算法，对待推荐项目进行过滤，生成相应推荐列表，并将最后的结果反馈给待推荐的用户。

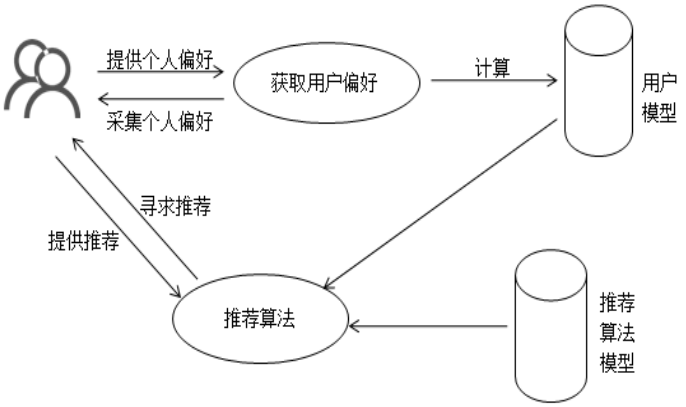

图 2-1 个性化推荐系统的通用模型

Adomavicius 等[15]人在 2005 发表的综述性文章中对推荐系统进行了公式化的描述：假设 U 表示所有的用户集合，S 表示所有待推荐的项目集合，在视频推荐领域代表的就是视频资源，f 表示效用函数，用于计算项目 s 对用户 u 的推荐度，或者可以理解成是用户 u 对项目 s 的兴趣度，则 f 可以描述成 $f: U \times S \rightarrow R$ ，其中 R 表示所有待推荐项目与用户之间的兴趣集合。推荐系统的基本思想就是找到使兴趣度 R 最大的那些推荐对象 S^*，如公式 2-1 所示：

$$\forall u \in U，\ S^* = arg\, max_{s \in S} f(u, s) \tag{2-1}$$

从以上的描述可以看出，一个完整的推荐系统主要包含三部分内容，分别是用户、推荐项目以及推荐算法。而其中最为核心内容的则是推荐算法，后文将对当前主流的推荐算法进行简要概述。

2.2 推荐系统的相关算法

如前文所述，在构成推荐系统的三项主要模块当中，最关键的是推荐算法，推荐算法的优劣，对整个推荐系统的好坏起着至关重要的影响。在推荐系统的研究领域中，推荐算法的研究也是最为热门的研究内容，大部分学者所发表的研究成果也都集中在这个方向。随着研究人员的不断探讨，涌现出了许多优秀的推荐算法。然而至今，对推荐算法还尚未有统一的归类准则，但目前主流的推荐算法主要可分为以下几种：基于内容的推荐算法[16]、协同过滤推荐算法[17]、基于关联规则的推荐算法[18]、基于社会化网络分析的推荐算法[19]、混合推荐算法等。下文将对以上各种推荐算法进行简单介绍并分析各自的优缺点。

2.2.1 基于内容的推荐算法

基于内容的推荐算法的思想最早来自于信息检索[21]和信息过滤[21]。它通过分析用户的喜好，并将与他喜好相似的项目推荐给用户。它通常是对项目本身的内容属性进行分析，提取出待推荐项目的内容特征，然后将提取到的内容特征与用户模型中偏好特征进行匹配，进而找到匹配度较高的待推荐项目推送给该用户。以图 2-2 为例，首先基于用户的历史行为数据对该用户所喜欢的视频的内容进行分析，发现该用户对武侠、爱情类的视频感兴趣，并将此作为用户的兴趣偏好进行建模。然后再分析他视频的内容特征，从视频 B 与视频 C 中选择出与用户兴趣偏好相似度高的视频 C 推荐给该用户。

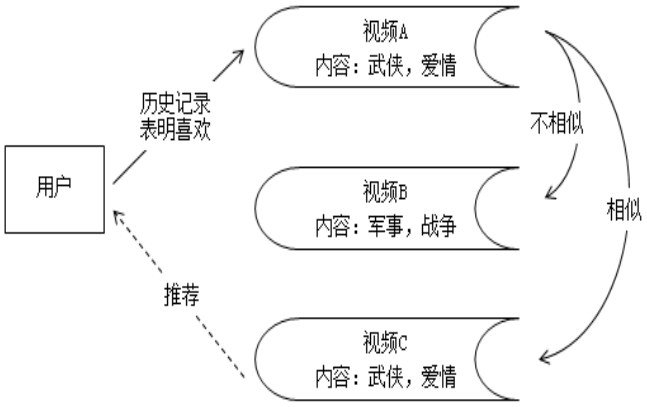

图 2-2 基于内容的推荐算法原理示意图

该算法通常是将每个对象描述为文本形式。然后利用相关技术对文本形式的内容进行处理和分析，构建用户的兴趣偏好模型。其中，最通常的方法是采用特征提取技术从非结构化的文本中抽取关键特征，建立一个特征向量来描述物品。基于内容的推荐算法通常会使用文本中的关键词作为该文本的特征表示。一个关键词 k_i 在文档 d_j 中的“重要性” $w_{i,j}$ 的描述手段有很多种，其中最著名的就是 TF-IDF（Term Frequency / Inverse Document Frequency）算法[22]，它认为一个词语的重要性随着它在文章中出现的频率增大而增加，而随着在语料库中出现的频率增大而减小。

TF-IDF 是这样定义的：假设所有文档的总数为 N，关键词 k_j 总共出现了 n_i 次，$f_{i,j}$ 表示为关键词 k_j 在文档 d_j 出现的次数，用 $TF_{i,j}$ 表示关键词 k_j 在文档 d_j 中出现的词频，公式为：

$$TF_{i,j} = \frac{f_{i,j}}{\sum_z f_{z,j}} \tag{2-2}$$

逆向文件频率 IDF_i 通常定义为：

$$IDF_i = log\frac{N}{n_i} \tag{2-3}$$

于是一个关键词 k_j 在文档 d_j 中的 TF-IDF 权重可以定义为：

$$w_{i,j} = TF_{i,j} \times IDF_i \tag{2-4}$$

于是一篇文档 d_j 可以通过获取的关键词权重进行描述，形式如下：

$$Content(d_j) = (w_{1,j} \ldots w_{k,j}) \tag{2-5}$$

上式就是基于 TF-IDF 算法对项目的特征建模。而对于用户而言，其向量特征则可

简单定义为成该用户所观看的所有文档的特征向量的加权平均。当完成对项目与用户的模型构建后，则可将待推荐对象的内容特征与用户的偏好特征进行匹配，并且计算出他们之间的的相似度[20]。而计算两个向量之间相似性的方法有很多种，在基于内容的推荐中常用到的是余弦相似度。当得到待推荐对象与用户之间的相似度后，将结果按照从大到小的顺序排列，并选出与用户相度最高的项目推荐给用户。

基于内容的推荐算法除了使用上面介绍的基于信息检索的方法外，还采用很多其他的方法。其中比较常见的有贝叶斯分类算法[23-24]，聚类分析算法，决策树算法，人工神经网络[25]等。这些算法与信息检索算法不同的是，它们是根据统计学和机器学习相关理论从已有的数据中通过分析得到用户的兴趣偏好模型，最后再根据模型匹配进行项目推荐。

Ricci 等[26]人给出了一个基于内容的推荐系统框架，如图 2-3 所示。

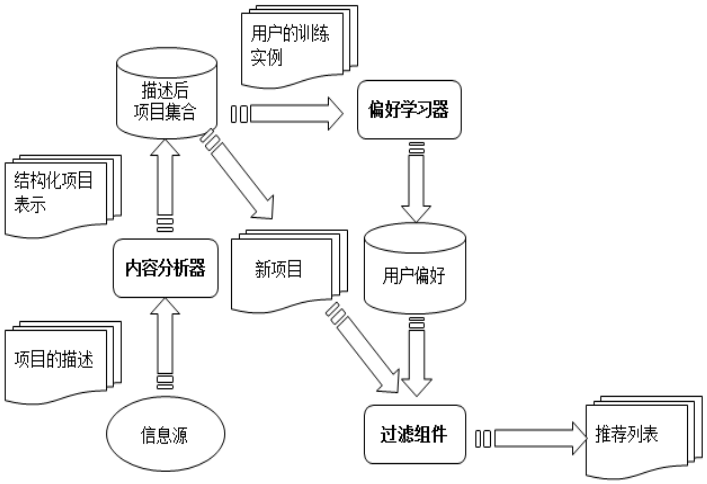

图 2-3 基于内容的推荐系统框架

根据图 2-3 中内容可以看出，基于内容的推荐系统通常包括以下几个步骤：

1）推荐项目特征建模。当有新的推荐项目加入时，系统都会按照相应算法对项目的内容进行分析，获取项目的特征，构建项目模型。

2）用户偏好获取，即用户对象建模。基于内容的推荐系统会根据用户的历史行为数据，分析用户的兴趣偏好，并安装一定的方式来构成用户的特征信息。除了之外，系统还可以让用户自己提供偏好信息，从而构建更全面的用户模型。

3）计算相似度。系统根据待推荐项目的内容特征和用户的兴趣偏好特征来计算两者之间的相似度。

4）项目过滤，生成推荐结果。将计算所得到的相似度依照递减的规律进行排序，

获取与用户最为相似的项目生成推荐列表，并将结果反馈给用户。

基于内容的推荐系统的优点在于：它简单有效，推荐结果符合人们的认知；它无须用到用户的历史评分信息，因而没有协同过滤中的数据稀疏性问题；没有新推荐项目的冷启动问题，即使刚加入的新项目也可以提取他的特征进行相似度计算。但该方法亦受存在其不足之处，受特征数量和特征类型的限制，需要较多的领域知识，如果不能得到足够的信息，则推荐效果较差；推荐结果较单一，基于内容的推荐系统是根据用户的评分历史进行推荐的，推荐的项目与用户已评分的项目过于相似，因此缺少了新颖性；它存在新用户的冷启动问题，当新用户出现时，由于没有用户的行为记录，因此就无法获取该用户的兴趣偏好，从而不能进行可靠的推荐。

2.2.2 基于协同过滤的推荐算法

协同过滤技术是在个性化推荐领域应用得最广泛且最成功的一项技术。它是一种基于一组兴趣相同的用户或项目进行的推荐，它根据与目标用户兴趣相似的用户的偏好信息产生对目标用户的推荐列表[27-29]。

基于协同过滤的推荐算法通常情况下可分为以下三类：基于用户的协同过滤推荐算法 (User-based Collaborative Filtering)、基于项目的协同过滤推荐算法(Item-based Collaborative Filtering)和基于模型的协同过滤推荐算法(Model-based Collaborative Filtering)。

1）基于用户的协同过滤推荐算法[30-31]

基于用户的协同过滤推荐算法的核心思想是寻找与目标用户兴趣偏好相似的用户，并将他们喜欢的项目推荐给目标用户。它不考虑用户、项目的属性信息，主要是根据用户对项目的偏好信息，发掘不同用户之间的品味相似性，并使用这种相似性来进行个性化推荐。它是基于这样的一个假设：如果一些用户对某一类项目的打分比较接近的时候，则我们在一定的程度上可以认为这些用户有相同的兴趣偏好，并以此作为依据认为他们对其他项目也会有相似的评分。图 2-4 给出了基于用户的协同过滤原理示意图。从图中可以看到，根据用户的历史行为记录，发现用户甲和用户丙同时对视频 A 和视频 C 感兴趣，而用户乙只对视频 B 感兴趣，则在某种程度上认为用户甲和用户丙具有喜好相似，于是便将用户甲看过的视频 D 推荐给用户丙。基于用户的协同过滤算法其核心思想就是通过相似性度量方法计算出最近邻居集合，并将最近邻的评分结果作为推荐预测结果返回给用户。

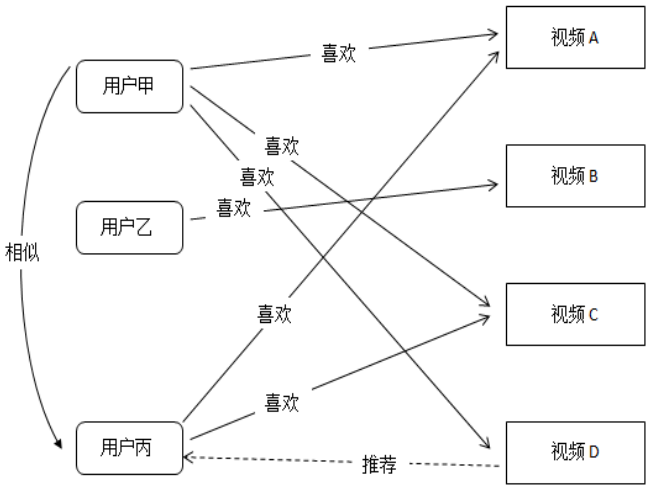

图 2-4 基于用户的协同过滤原理示意图

通常情况下，基于用户的协同过滤推荐算法可以归纳为以下三步：1）根据用户的历史行为数据获取每个用户兴趣偏好特征；2）是依靠所获取的兴趣偏好特征，按照一定的方式计算用户之间的相似性，并以此作为推荐的依据目标用户的喜好进行预测；3）生成推荐结果，采用计算相似度的方法获得具有类似喜好或兴趣的相邻用户，然后按相似度的大小进行排序，选择若干个邻居所喜欢的项目来做推荐。具体实现方式如下[32]：

（1）收集每个用户的兴趣信息。根据用户的历史行为数据构建一个$n \times m$的评分矩阵。矩阵中的数据通常有三种类型：第一种布尔类型，例如是否观看某电影，是否购买某商品；第二种数值类型，例如用户的评分数据；第三种非可量化数据，例如用户的评论等。然后根据该矩阵的信息通过一定方法将用户表述成向量形式。

（2）计算用户的最近邻集。这是基于用户的协同过滤算法中最为关键的步骤。当把用户表示成向量的形式后，我们便可以通过计算两个向量之间的距离来描述用户之间相似性。常见的几种计算向量距离的方法有：

A、余弦相似度

余弦相似度的大小与向量的长度无关，它表示的是两个向量之间的夹角大小，他在一定程度上可以反应出两个人的兴趣爱好方向是否一致。它的表达式为：

$$cos(U_1,\ U_2) = \frac{\sum_{i=1}^{K} u_{i,1} \cdot u_{i,2}}{\sqrt{\sum_{i=1}^{K} u_{i,1}^2} \cdot \sqrt{\sum_{i=1}^{K} u_{i,2}^2}} \tag{2-6}$$

B、Peason 相关性系数

$$Peason(U_1,\ U_2) = \frac{\sum_{i=1}^{K} (u_{i,1} - \overline{u_1}) \cdot (u_{i,2} - \overline{u_2})}{\sqrt{\sum_{i=1}^{K} (u_{i,1} - \overline{u_1})^2} \cdot \sqrt{\sum_{i=1}^{K} (u_{i,2} - \overline{u_2})^2}} \tag{2-7}$$

C、明可夫斯基距离

$$Dist(U_1,\ U_2) = (\sum_{i=1}^{k}|u_{i,1}\ - u_{i,2}|^{p})^{\frac{1}{p}} \quad (2\text{-}8)$$

当 p=2 时，就是我们常见的欧式距离。

（3）生成推荐结果。按照上一步中得到的相似度大小，生成一组用户的最近邻集，最后然后把该集合中用户所喜好的项目推送给目标用户完成推荐。

2）基于项目的协同过滤推荐算法[33-35]

基于项目的协同过滤算法是寻找相似的推荐项目，通过用户对相似项目的评分数据预测目标项目的评分。它同样不考虑用户和项目的属性信息，它根据用户对项目的偏好信息，发掘不同项目之间的相似性。该算法的基本基本思想是认为，如果大部分用户对某些项目的打分比较接近，则当前用户对这些项目的打分也会比较接近。

图 2-5 给出了基于项目的协同过滤示意图，从图中我们不难看出，视频 A 与视频 C 都同时被用户甲和用户乙喜欢，而视频 B 只有用户乙喜欢，根据基于项目的协同过滤的思想认为视频 A 和视频 C 在相当程度上具备一定的相似性，于是当用户丙对视频 A 感兴趣时，则认为用户丙也会对与视频 A 相似的视频 C 感兴趣，并将视频 C 推荐给用户丙。

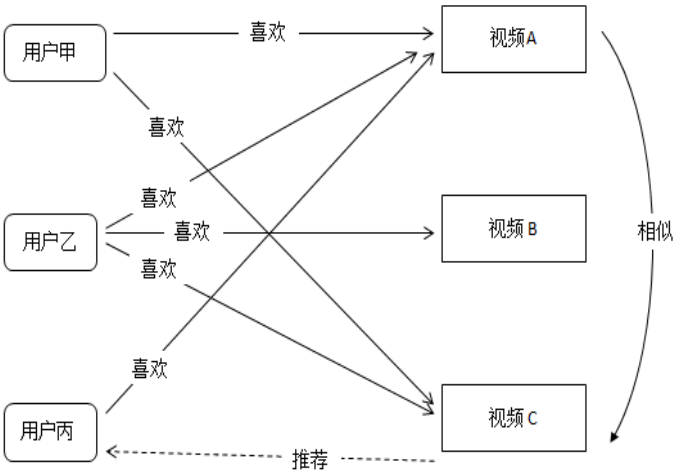

图 2-5 基于项目的协同过滤原理示意图

基于项目的协同过滤的推荐过程也大致分为三步：

（1）收集用户偏好信息。这一步与基于用户的做法大致相当，此处就不再赘述。

（2）计算项目的最近邻集。此处计算项目相似性的方法也与基于用户的协同过滤算法相同，包括余弦相似度，Peason 距离以及明可夫斯基距离等。

（3）生成推荐结果。根据第二步中计算得到的结果，按照从大到小的顺序进行排

列，将与目标项目相似度较高的 Top-N 个项目推荐给目标用户。

3）基于模型的协同过滤推荐算法

基于模型的推荐算法的大致思想是通过机器学习和统计的方法通过离线计算实现推荐。它通常会根据用户的历史行为数据，对用户的偏好模型进行训练，然后在此模型基础上对用户的兴趣进行预测，最后生成推荐结果反馈给用户。在这个方法中建立用户模型是核心，常用的方法有贝叶斯网络[36]、基于奇异值分解[37]、基于决策树模型、聚类模型[38]等。

与基于内容的推荐算法相比，基于协同过滤的推荐算法拥有以下的优点：它能处理复杂的非结构化数据，如电影、图像等；能够发现用户的新爱好，丰富推荐的多样性；通过协作的方式分析用户之间的喜好，避免在基于内容的推荐中遇到特征提取不完全的情况。

然而，基于协同过滤的推荐算法也存在着以下缺点：它存在冷启动问题。新用户由于得不到他的兴趣偏好无法推荐，而新项目也因为没有用户评价信息也得不到推荐；存在数据稀疏性问题。在推荐系统中，相比需要预测的评分数据，已知的评分数据通常是极少的，从少量已知数据中获得有效评分预测非常重要。

2.2.3 基于关联规则的推荐算法[39]

基于关联规则的推荐算法起源于数据挖掘领域，主要是根据统计学手段，构建关联规则，以此作为依据完成推荐工作。说到基于关联规则，就不得不提到最著名的“啤酒喝尿布”的故事：在德国，人们发现许多父亲在超市为自己的孩子购物买尿布的时候，也会偷偷地为自己买一箱啤酒作为奖励，因此根据这一情况，超市的销售人员将啤酒喝尿布摆放在了一起，最后发现尿布和啤酒销量得到了巨大提高。基于关联规则的主要思想是挖掘物品与物品之间的相关性大小。

基于关联规则的推荐算法主要流程为：首先在数据结合中生成频繁项集，即计算各个推荐项目的高频组合；然后基于上述高频组合生成相应关联规则；最后根据该规则将相关项目推荐给目标用户。

基于关联规则的推荐算法其核心的关联规则的构建，人们常用到的关联规则算法包括有 Agrawal 等[40]人提出的 Apriori 算法和 Han 等[41]人提出的 FP-groth 算法。

2.2.4 基于社会化网络分析的推荐算法[42]

基于社会化网络分析的推荐算法旨在详尽地描述用户与用户之间的社会关系网络

[43]。我们常见的一些社会关系有双边关系，例如亲戚、朋友、同学等；单边关系，例如微博上的单向关注，个人微信公众号的单方订阅等。该算法将社会化的网络关系以图的形式展现，探索与分析各节点、边的重要程度，利用这些重要关系来进行推荐。举个简单例子：假设我们通过社会化网络构建，了解到用户 A 和用户 B 是互为同学关系。当用户 A 在网上购买了一本习题册后，我们便在一定程度上认为用户 B 也对该习题册感兴趣，从而将其推荐给 B。基于社会化网络推荐算法其核心思想是找到与目标用户社会关系相近的用户群体，然后将他们的喜好的物品推送给目标用户。

2.2.5 混合推荐算法

从上面的介绍我们可以看出单一的推荐算法或多或少都会存在各自的缺陷，为了尽可能的减少这些缺陷带来影响，人们提出了混合推荐策略，即融合不同的推荐算法进行推荐。在学术界和工业界提出了多种混合推荐系统[44-46]，理论上可以有很多种推荐组合方式，但最常见的组合是结合基于内容和基于协同过滤的推荐算法。Adomavicius 等人根据处理方式的不同将混合方法分为了以下几个类别：

1）并行结合的混合系统。即先分别实现协同过滤推荐系统和基于内容的推荐系统，然后再对它们的结果进行处理。其中的一种处理方法是直接对两种结果进行线性的加权权平均；另外一种处理方法是只推荐当前条件下性能指标更好的结果。

2）将基于内容的推荐算法融合到协同过滤推荐算法中[44]。

3）将协同过滤推荐算法融合到基于内容的推荐算法中[45]。

2.3 推荐系统的评价指标

通常情况下，一个新的推荐系统投入到实际应用之前，人们需要对它的性能进行评估。比较常见的推荐系统评价指标有：精确度、准确率、召回率、覆盖率和新颖性等。

2.3.1 预测精确度

在以预测评分为主的推荐系统中，通常是计算预测评分和真实评分的差异来表示系统的精确度。其中最为经典的就是平均绝对误差（Mean Absolute Error,简称 MAE），如果用 p_{ia} 和 r_{ia} 分别表示推荐系统所预测用户 i 对项目 a 的评分和用户 i 对项目 a 的真实评分， T 表示测试集数据。则 MAE 的计算形式如下：

$$MAE = \frac{1}{|T|}\sum_{i,a\in T}|p_{ia} - r_{ia}| \tag{2-9}$$

除此此外，还经常有用到均方根误差（RMSE）和标准平均误差（NMAE），它们

的公式分别如下：

$$RMSE = \sqrt{\frac{1}{|T|}\sum_{i,a\in T}(p_{ia} - r_{ia})^2} \tag{2-10}$$

$$NMAE = \frac{MAE}{r_{max} - r_{min}} \tag{2-11}$$

其中，r_{max} 和 r_{min} 分别表示用户评分的最大值和最小值。

而在以 Top-N 推荐为主的推荐系统中，它的目标是通过分析用户的历史日志信息找到用户未使用过的项目中用户兴趣度预测值最高的前 N 个项目推荐给该用户。因此它常用衡量推荐准确度的指标可以用准确率（Precision）和召回率（Recall）。

准确率表示的是推荐系统所推荐的项目中用户真正喜欢的项目的比例，计算公式如下：

$$Precision = \frac{N_{rs}}{N_s} \tag{2-12}$$

召回率表示的是在用户所有喜欢的项目中被系统推荐给用户的所占比例，计算公式如下：

$$Recall = \frac{N_{rs}}{N_r} \tag{2-13}$$

其中，N_{rs}表示推荐列表中用户喜欢的项目的个数，N_s 表示所有被推荐的项目的个数，N_r表示用户喜欢的所有的产品的个数。

2.3.2 ROC 曲线和 AUC 值

常用于评价一个分类器的性能好坏，曲线越饱满则说明分类效果越好。而 AUC 是指曲线下方的面积大小，因此它的值越大则说明分类器的效果越好。具体来说，ROC 曲线是根据下面两个指标得到的：

$$TPR = \frac{TP}{TP+FN} \tag{2-14}$$

$$FPR = \frac{FP}{FP+TN} \tag{2-15}$$

其中，TPR（true positive rate）是指真实正样本比，表示分类器对所有正样本预测正确的比例大小。FPR（false positive rate）是指误判正样本比，表示分类器对所有负样本预测错误的比例大小。

2.3.3 准确度（Accuracy）

该指标用于整体上最分类器的性能进行评估，在本文它反映了系统对用户点击行为预测的准确度，它计算公式为：

$$acc = \frac{TP+TN}{TP+TN+FP+FN} \tag{2-16}$$

除此之外，在推荐系统中还会经常用到覆盖率和热门度。覆盖率用于描述推荐系统对冷门物品的推荐能力，它的意义在于某些情况下用户可能也会对冷门的项目感兴趣，而且提高覆盖率也可以对系统中的项目资源进行充分利用。而热门度则是用来描述一个项目同时被多少人喜欢，它可以用于评价一个系统的新颖性

2.4 本章小结

本章主要是对个性化推荐系统的理论知识进行相关介绍。首先介绍了推荐系统的基本原理，并给出了推荐系统的基本框架；然后对推荐系统中的主流的推荐算法进行了阐述，并总结了各自的优缺点；最后对推荐系统中评价指标进行了简单介绍。

第 3 章　基于深度语义模型的视频个性化推荐系统

在视频推荐领域，多数情况下人们会选择基于协同过滤的推荐算法进行推荐工作，从往年的论文发表情况来看也可以证实这一点。相较与基于内容的推荐算法人们之所以会倾向于协同过滤，究其根本原因在于网络视频的媒资信息参差不齐，内容繁杂，人们难以采取有效的手段来对视频的内容特征进行提取，进而对视频内容的进行分析。

但是近些年来，随着互联网行业和多媒体技术的飞速发展，网络视频量以及视频用户量都在迅速增长。据统计，目前腾讯视频的视频总量达到 1500 万，月度活跃用户更是高达 2 亿，每天有接近 10 亿的视频播放量，日播记录数据拥有数百 GB。如此大规模的视频以及用户总数必然会使得协同过滤中的评分矩阵变得异常稀疏，极大地限制了协同过滤的推荐效果。与此同时深度学习[47]在自然语言处理、图像处理和语音处理等领域取得了技术突破，为人们对视频内容的特征提取提供一定的技术支持，使得人们看到了在视频推荐领域的使用基于内容的推荐算法的希望。

本章提出了一种基于深度语义模型的视频的个性化推荐算法。该算法利用了深度学习在自然语言处理上的相关技术，分析并处理视频的文本内容信息，提取其语义层面的特征向量，构建用户以及视频模型，深度挖掘用户与视频之间的内在关系，最后进行视频的个性化推荐。

3.1 算法流程概述

图 3-1 给出了算法的整体流程示意图，从图中可以看到，基于深度语义模型的视频个性化推荐算法大致包括以下几个部分：

1）数据采集与预处理。通过 Hadoop 大数据平台获取后台数据库中的用户的历史行为数据以及视频的文本描述数据，然后对数据进行预处理；

2）用户和视频建模。利用文本挖掘的相关算法对视频的文本数据进行处理，提取视频内容的相关特征，然后再根据用户的历史点击行为记录，来完成用户的特征描述；

3）构建并训练深度神经网络模型。根据用户的历史点击行为记录建立用户-视频点击对作为神经网络的训练样本，并以用户和视频的文本特征向量作为神经网络输入，对网络进行有监督的训练。利用深度学习的特征提取能力，获取视频与用户的语义特征，

深度挖掘用户与视频之间潜在关系，完成模型构建；

4）生成推荐项目。利用训练好的模型，计算用户与视频之间的相似性大小，以此作为依据生成推荐列表并将结果反馈给待推荐用户。

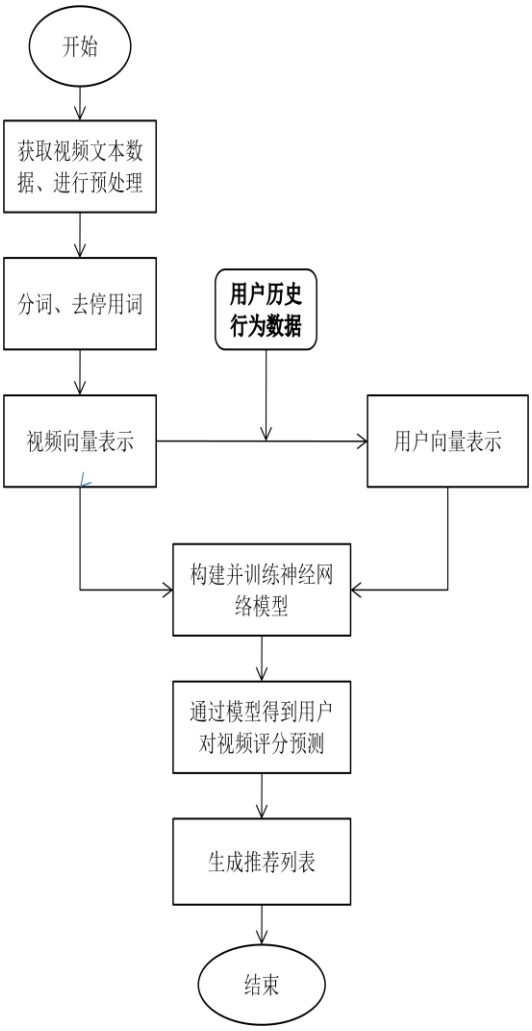

图 3-1 推荐算法流程示意图

3.2 数据获取与数据预处理

3.2.1 数据获取

在整个推荐算法的流程中首先需要获取相关数据，文中所用到的数据主要包括两方面内容：

第一部分是视频的文本描述信息，一个典型的视频文本描述信息的具体内容包含有：视频的标题、视频的内容简介、视频的制作地区、视频的导演、主演和编剧、视频的标签信息以及视频的上传时间和上传作者等。这部分内容主要用于视频内容特征提取，构建视频的模型。

第二部分是用户对推荐视频的历史点击行为记录。当然这里的点击行为并非就是指用户打开视频链接即可，某些情况下用户点击视频只是粗略的浏览视频内容，而并非代表用户对该视频真正感兴趣。因此为了使得用户的点击数据有意义，在判断用户的一个点击行为是否有效时本文加入播放时长比这一判决条件，具体公式如下：

$$\tau_{ia} = \frac{t_{ia}}{T_i} \tag{3-1}$$

其中，t_{ia} 表示用户a观看视频 i的时长，T_i表示视频i的总时长。只有当播放时长比 τ_{ia}大于某个阈值时，我们才认为用户的这个点击行为是有效的，并对其进行记录，否则都视为无效操作。在实际操作中，我们将阈值设为 0.3。这部分数据用于分析用户的兴趣偏好，构建用户模型。

3.2.2 数据预处理

通常情况下，我们拿到的视频媒资信息是存在有各种各样的问题的，例如：数据缺失、数据冗余等。为了使我们的数据在推荐过程中具有完整性和可靠性，保证最后的推荐结果。在正式处理分析之前，我们需要对原始数据进行预处理，这里主要包括视频的过滤和视频的融合处理。视频的过滤主要是针对上述视频文本信息字段缺失的视频，这类视频大部分来自于用户自己制作的视频集合中，也就是熟称的 UGC 类别的视频。而视频的融合处理主要是针对视频的大部分字段信息相同的视频，这类视频主要来自于腾讯视频官方合作伙伴所提供动漫类和电视剧类聚集视频。这类视频数据结构标准化，描述信息大同小异，因此可以进行融合处理。

3.3 视频与用户建模

通过上述的预处理环节，完成数据的清洗之后，我们便利用自然语言处理的相关知识对视频与用户分别进行模型构建。

3.3.1 文本分词

因为本系统处理的是中文视频的文本信息，所以我们第一步需要对文本内容进行中文分词处理。在本系统中，我们所使用的分词算法是中国科学院技术研究所张华平[48]等人研究与开发的 ICTCLAS 分词系统。该分词系统分词速度快，可达 996kb/s，分词准确率高，精度高达 98.45%，并且还支持用户自带词典，即使对较新网络用词也能取得一个不错的分词效果。除此之外，ICTCLAS 还能对某些专有名词进行很好的识别，具备去停用词的功能。

3.3.2 视频建模

近些年来，越来越多的视频服务的供应商开始使用标签来描述视频的基本信息。因为它能够用简短精悍的字词描述搭建起沟通用户兴趣与物品语义的桥梁[49]。它是对物品一个行之有效的描述，不管对用户兴趣建模还是物品特征提取都是具有重要价值的原始材料。本文通过提取视频文本描述信息中的较为重要的关键词作为视频的标签信息，并以此来构建视频模型。

当通过前面的分词算法完成对视频文本内容的分词处理之后，本文使用自然语言处理中经常用到的 TF-IDF 算法来进行关键词提取。该算法的具体计算公式在本文的第二章中有详细介绍，这里就不再赘述。它的大体思想是认为：一个词在一篇文档中的重要性取决于两个方面：一个方面就是该词语在该文档中出现的频率，被称作单词词频，也就是所谓的 TF 值。它在某种程度上直接体现了词语的在该文章中的重要程度，词语出现的频率越高，它的重要性越高，TF 值就越大；另一个方面就是这个词语在所有文档库中的出现的频率，被称作逆向文件频率，这就是所谓的 IDF 值。它反映这个词语在整个文档库中的重要程度，这是因为一个关键词如果在所有文档中经常出现，那么用它来代表某篇文档的能力反而会降低。从信息论的角度来看，是因为这类词语所提供的信息量较少。因此当该频率越高时，IDF 的值就应该越小。最后 TF-IDF 算法就是在这两个频率的基础上进行关键词的权重计算的。

当对所有的视频的文本内容使用 TF-IDF 算法之后，就能够得到各个关键词在各自视频中重要程度的量化数值。然后过滤掉权重较轻的关键词，保留那些对视频较为重要的关键词作为该视频的标签信息，用于描述视频的特征，以此构建视频基本模型。除此之外，我们还对整个视频的标签信息进行统计，构建视频标签词库，统计标签总数，建立标签索引。

于是，一个视频的特征信息便可以描述成如下的向量形式：

$$Vedio = \{t_1,\ t_2 \dots t_n\} \tag{3-2}$$

其中，t_i 表示该视频在第 i 个标签上的 TF-IDF 权重，n 表示标签的总数。

3.3.3 用户建模

用户模型的建立是基于上述视频向量而得到的。首先，根据用户的历史点击行为记录，统计用户过去所看过的所有视频的信息。前文已经有说明，只有当用户的点击行为有效时，才认为该用户对该视频感兴趣。然后，便将统计得到的所有视频向量进行加权

平均以此作为用户向量模型，用于描述用户的特征信息，构建用户模型。

具体而言，假设某用户 u 在过去的一个月时间里，一共点击了 n 部视频，则描述该用户特征向量的计算公式如下：

$$User = \frac{1}{n}\sum_{i=1}^{n} V_i \tag{3-3}$$

其中，V_i 表示视频 i 的向量，具体形式如公式 3-2 所示。

3.4 深度语义模型搭建与训练

当完成视频建模以及用户建模后，接下来的工作便是要设计方法计算用户与视频之间的相似度大小，以此作为视频推荐的评分依据。最直接的方法就是拿上述方法中得到用户向量和视频向量通过第二章中提到的相似性的计算方法进行相似度的计算。但这种方法存在着较大的问题。因为，从上面介绍的用户和视频模型构建的方法可来看，用户和视频特征其实可以理解成仅仅是由一组带有权重的标签进行描述的。然而各个标签之间实际上独立，并没有分析和挖掘标签与标签之间的内在联系。例如：对于标签“中国”和“美国”而言，它们之间实际上具有一定程度的相关性，而并非相对独立。对于某个喜欢看时事新闻视频的用户而言，其实无论是中国或者美国的实时报道都非常感兴趣，如果仅利用上面提到的视频和用户向量信息进行相似度计算，中国实时视频和美国实时视频对他而言可能会出现较大偏差。因此直接计算的结果必然会有很大的局限性。

为此构建了一个深度神经网络模型。该网络将用户的历史点击行为记录作为训练样本，以用户和视频的特征向量作为网络输入，挖掘与分析用户与视频之间的深层次的语义关系。该模型通过利用深度学习的特征提取能力，对视频和用户的低层特征信息进行组合和提取，生成更加抽象高层特征，用于对用户和视频数据进行分布式表示，并在此基础上计算用户与视频的相似性，进而进行用户群的过滤，生成推荐结果。

3.4.1 深度模型框架

在本文中所提出的深度神经网络模型框架如图 3-2 所示。从图中可以看到，该网络模型整体主要由两部分组成，包括一个用户模型和一个视频模型。它的输入是一组高维空间向量，也就是前文中计算所得到的用户和视频的标签向量，该向量可以看成是视频和用户的文本内容的直接表述。它的输出则是一组低维稠密的语义向量空间，它可以看成是神经网络对低层文本特征进行特征提取后的抽象化描述。

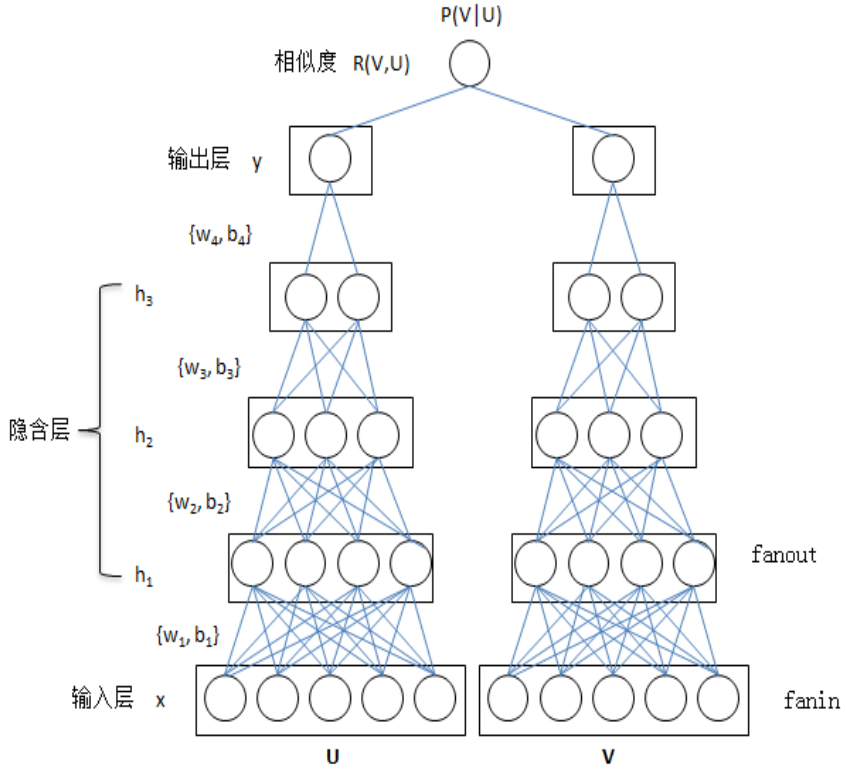

图 3-2 深度神经网络模型示意图

由上图可知本推荐策略的大致流程是：首先将视频和用户的文本特征分别映射成两个高维的空间向量，并以此作为神经网络的输入。然后根据用户的点击行为记录作为训练样本对网络模型参数进行训练，并利用神经网络的对视频及用户进行特征提取，将视频和用户的原始特征信息进行逐层的抽象，并映射成两个低维的语义相关的特征向量。最后根据余弦距离来计算用户与视频的相关性大小，以此作为推荐的评分依据。

此网络的具体描述如下：假设用 x 表示输入向量，y 表示输出向量，h_i 表示神经网络中的隐含层，$i = 1，2 \ldots N-1$，W_i 表示网络中第 i 层的权重矩阵，b_i 表示第 i 层的偏置。则有以下公式：

$$h_1 = W_1 x + b_1 \tag{3-4}$$

$$h_i = f(W_i h_{i-1} + b_i), i = 2, \ldots, N-1 \tag{3-5}$$

$$y = f(W_N h_{N-1} + b_N) \tag{3-6}$$

其中，$f(x)$ 表示激活函数，在本文中用$tanh$作为隐藏层和输出层的激活函数。其具体公式如下：

$$f(x) = \frac{1-e^{2x}}{1+e^{2x}} \tag{3-7}$$

最后计算用户 U 和视频 V 的语义相关性大小 $R(U,V)$ ，计算式子如下：

$$R(U,V) = cosine(y_U, y_V) = \frac{y_U{}^T y_V}{\|y_U\|\|y_V\|} \tag{3-8}$$

其中，y_U 和 y_V 分别表示经过语义特征提取后的用户和视频的分布式向量表示。在整个推荐过程，当得到最后视频与用户的相关性评分后，将该评分按照从大到小的顺序进行排列，并取 Top-N 作为最终的推荐结果，反馈给用户。

3.4.2 网络模型理论推导

在机器学习领域中，有一种叫做判别模型的建模方法。该方法是对未知数 y 和已知数据 x 之间的关系进行建模。它是一种基于概率论的方法，当已知输入变量 x 后，判别模型构建条件概率分布 $P(y|x)$ 来对 y 进行预测。

受到判别模型的启发，我们对整个推荐问题进行变换思考，可以将其看成在已知用户兴趣偏好的条件下，用户观看某视频的条件概率的大小。不难理解，用户与视频之间是存在着一定程度关联的，已有的历史行为记录可以看成是该用户的一个先验知识。基于这一思想，本文中设计了一种监督式的训练方式，用于对整个网络模型的参数进行学习和训练，包括网络中的权重矩阵 W_i 和偏置向量 b_i。具体的推导过程如下：

首先，我们知道一个用户对视频的感兴趣程度可以通过他们的语义相关性的大小来反应，即 $R(U,V)$ 的大小。然后对于一个特定的用户 U 而言，他点击观看某个视频 V 的概率的大小，可以用条件概率 $P(V|U)$ 来表示，并将它定义为：

$$P(V|U) = \frac{exp(\gamma R(U,V))}{\sum_{V' \in \boldsymbol{V}} exp(\gamma R(U,V'))} \tag{3-9}$$

其中，γ是平滑因子， $\boldsymbol{V}$ 表示一组候选视频集合。理论上，集合 $\boldsymbol{V}$ 应该是包括所有的候选视频的集合，即所有的正样本和负样本。然而在实际操作中，负样本是非常大的，为了节约计算成本，提高模型的训练效率，一般会对负样本按照一定的方式进行负采样。在本算法中，使用的采样方式为：取一个正样本组合(U,V^+) 和三个个负样本(U,V^-)组合作为候选视频集合 $\boldsymbol{V}$ 。在实验过程中发现，样本的组合对最终的结果影响不大。于是可将公式 3-9 可以近似地描述成：

$$\mathrm{P(V^+|U)} = \frac{exp(\gamma R(U,V^+))}{\sum_{V' \in \boldsymbol{V}} exp(\gamma R(U,V'))} \cong \sigma(U,V^+) \cdot \prod_{j=1}^{3} \sigma(U,V_j^-) \tag{3-10}$$

其中，$\sigma(U,V^+) = \frac{1}{1+exp(\gamma R(U,V^+))}$ 。

在确定条件概率的形式后，我们可以用最大似然函数[50]估计来对该神经网络中的参

数进行训练，即等价于对负对数似然函数进行最小化计算。负对数似然函数构造如下所示：

$$L(\theta) = -log \prod_{(U,V^+)} P(V^+|U) = -log \prod_{(U,V^+)} (\sigma(U,V^+) \cdot \prod_{j=1}^{3} \sigma(U,V_j^-))$$
$$= -\sum_{(U,V^+)}(\ log\sigma(U,V^+) + \sum_{\mathrm{j}} log\big(1-\sigma(U,V^-)\big)) \quad (3\text{-}11)$$

上式便是本神经网络的中最终需要优化的目标函数。其中 θ 表示的是神经网络中的参数集合 $\{W_i, b_i\}$。

为了统一表示，设定每个正样本的 label 为 1，每个负样本的 label 为 0，则每个样本的负对数似然函数就可以表示成下面的形式：

$$f = -label \cdot log\sigma(U,V) - (1-label) \cdot log\sigma(U,V) \quad (3\text{-}12)$$

由于 $L(\theta)$ 对于参数 θ 是可导的，于是便可以使用梯度下降法对目标函数进行优化，参数的更新法则为：

$$\theta_t = \theta_{t-1} - \lambda_t \frac{\partial L(\theta)}{\partial \theta}|_{\theta=\theta_{t-1}} \quad (3\text{-}13)$$

其中，λ_t 是第 t 次迭代的学习率，θ_t 和 θ_{t-1} 分别为网络模型在第 t 次和第 $t-1$ 次迭代的网络模型参数。

下面将对偏导数 $\frac{\partial L(\theta)}{\partial \theta}$，即神经网络的中参数的梯度的计算过程进行详细推导。

3.4.3 求导过程推导

假设训练样本的总数为 N，使用 $\left(U_i, V_i^+\right)$ 表示第i个正样本。则每个正样本的负对数似然函数如下：

$$L_i(\theta) = -logP\left(V_i^+ \middle| U_i\right) \quad (3\text{-}14)$$

于是整个目标函数的偏导数可表示为：

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{i=1}^{N} \frac{\partial L_i(\theta)}{\partial \theta} \quad (3\text{-}15)$$

因此只需要求得单个样本目标函数的偏导数，便可以得到整个样本的偏导。接下来主要是推导单个样本的偏导求解过程。

对于每一个用户 U 和每一个视频 V，用符号 $h_{i,U}$ 和符号 $h_{i,V}$ 分别表示视频和用户网络中第 i 个隐含层的激活值，$\mathrm{y_U}$ 和 $\mathrm{y_D}$ 则表示输出层的激活值，其具体的计算公式见 3-4、3-5、3-6。

将公式 3-9 带入到 3-14 中，为了简化起见，此处已把下标 i 去掉，于是得到：

$$L(\theta) = log\ (1 + \sum_j exp\ (-\gamma \Delta_j)) \tag{3-16}$$

其中 $\Delta_j = R(U, V^+) - R(U, V_j^-)$ 。

于是目标函数对输出层的权重矩阵求偏导得到结果如下所示：

$$\frac{\partial L(\theta)}{\partial W_N} = \sum_j \alpha_j \frac{\partial \Delta_j}{\partial W_N} \tag{3-17}$$

其中：

$$\frac{\partial \Delta_j}{\partial W_N} = \frac{\partial R(U,V^+)}{\partial W_N} - \frac{\partial R(U,V_j^-)}{\partial W_N} \tag{3-18}$$

$$\alpha_j = \frac{-\gamma exp\ (-\gamma \Delta_j)}{1+\sum_{j'} exp(-\gamma \Delta_{j'})} \tag{3-19}$$

为了简化起见，这里分别使用a、b、c表示 ${y_U}^T y_V$ 、$\frac{1}{\|y_U\|}$、$\frac{1}{\|y_V\|}$。则 3-16 式可以变换成如下形式：

$$\frac{\partial R(U,V)}{\partial W_N} = \frac{\partial}{\partial W_N} \frac{{y_U}^T y_V}{\|y_U\|\|y_V\|} = \delta_{y_U}^{(U,V)} h_{N-1,U}^T + \delta_{y_V}^{(U,V)} h_{N-1,V}^T \tag{3-20}$$

其中：

$$\delta_{y_U}^{(U,V)} = (1 - y_U)^\circ(1 + y_U)^\circ(\mathrm{b}c y_V - acb^3 y_U) \tag{3-21}$$

$$\delta_{y_V}^{(U,V)} = (1 - y_V)^\circ(1 + y_V)^\circ(bc y_U - abc^3 y_V) \tag{3-22}$$

其中，$^\circ$ 表示 element-wise 乘法。

对于隐藏层权重的更新，我们则使用误差反向传播算法进行梯度的计算。对于第 i 个隐含层其残差的计算公式为：

$$\delta_{i,U}^{(U,V)} = \left(1 + h_{i,U}\right)^\circ\left(1 - h_{i,U}\right)^\circ W_{i+1}^T \delta_{i+1,U}^{(U,V)} \tag{3-23}$$

$$\delta_{i,V}^{(U,V)} = \left(1 + h_{i,V}\right)^\circ\left(1 - h_{i,V}\right)^\circ W_{i+1}^T \delta_{i+1,V}^{(U,V)} \tag{3-24}$$

其中，$\delta_{N,U}^{(U,V)} = \delta_{y_U}^{(U,V)}$和$\delta_{N,V}^{(U,V)} = \delta_{y_V}^{(U,V)}$。

于是对于神经网络中隐含层的权重矩阵 $W_i, i = 1,2,\dots,N-1$ 的计算公式可以表示为：

$$\frac{\partial L(\theta)}{\partial W_i} = \sum_j \alpha_j \frac{\partial \Delta_j}{\partial W_i} \tag{3-25}$$

其中：

$$\frac{\partial \Delta_j}{\partial W_i} = \left(\delta_{i,U}^{(U,V^+)} h_{i-1,U}^T + \delta_{i,V^+}^{(U,V^+)} h_{i-1,V^+}^T\right) - \left(\delta_{i,U}^{\left(U,V_j^-\right)} h_{i-1,U}^T + \delta_{i,V_j^-}^{\left(U,V_j^-\right)} h_{i-1,V_j^-}^T\right) \quad (3\text{-}26)$$

以上便是整个网络中权重矩阵 W_i 的更新的推导过程，对于偏置 b_i 的推导和权重矩阵的相似，本文就不再赘述。

3.5 生成推荐列表

当完成对神经网络模型的训练之后，便可将目标视频与目标用户的向量信息输入到神经网络模型当中，最后通过训练好的神经网络模型得到用户对该视频的预测评分值。以此为依据，将该值按大小顺序进行排列，取评分较高的前 N 个视频生成推荐列表，推送给用户，完成推荐。

3.6 实验设计及结果分析

为了说明本章所提出的基于深度语义模型的视频推荐算法的有效性，本文从实践的角度设计了多组对比实验。下文将详细介绍实验的设计过程以及实验结果。

3.6.1 实验数据及实验环境

本实验所使用的实验数据来自于某视频网站，主要包括两部分内容：一部分是 15 天所有视频的文本内容描述信息，摘取部分示例如表 3-1 所示；另一部分是 15 天中所有用户对推送视频的历史点击行为数据，数据形式如表 3-2 所示。对于表 3-1 和 3-2 中的视频 ID 和用户 ID 是经过脱敏处理后的唯一标识。

表 3-1 视频文本描述数据示例

视频 ID	标题	视频内容简介
video135635	高离婚时代要不要婚前财产上“保险”	近几年，随着经济的发展，人们的生活环境发生了变化，思想观念上也出现了变化，中国居民的离婚率呈直线上升趋势。中国进入了高离婚时代。在这个时代下，需不需要给婚前财产上一份保险呢？
video100439	高圆圆韩国行美出新高度，现场直击女神时髦换装	在清潭洞小街小巷，我们用行走和享乐解读高圆圆当下的状态。她总是称自己是小镇女孩，随性，自由，在北京独自骑车出门，逛超市，从不刻意的去武装自己，“太过于小心翼翼，太辛苦了，人嘛需享乐！不要太在意他人的说法”看过这期《报尚名来》你会更加的喜欢高圆圆！

video102149	高校校长向毕业生鞠躬道歉：对不起，食堂饭菜不好吃	朴实的话语，诚挚的道歉，深深的鞠躬。6 月 29 日上午，在安庆师范大学毕业典礼暨学位授予仪式上，校长闵永新的这一“意外之举”，让在场学生和老师深受感动。
video153404	让孩子独立自强而非溺爱，这才是最好的家庭教育	授之鱼不如授之以渔，家庭教育真的是一门学问，父母是孩子的第一任老师，言行举止都影响着孩子的一生，看泰国妈妈如何教会自己女儿受用一生的事情。

表 3-2 用户对推送视频的历史点击行为记录示例

用户 ID	视频 ID	是否点击	时间
user234234	video183489	0	2016-06-07 19:28:01
user648999	video723431	1	2016-06-07 17:38:01
user120948	video123932	0	2016-06-07 09:18:40
user234944	Video098901	1	2016-06-07 13:17:53
user599821	video116341	0	2016-06-07 21:30:40

本实验总共获取了数亿条的用户的历史点击行为记录数据，其中正样本约为 1000 万条，按照 3:1 的比例对负样本进行负采样，最后得到总共 4000 万条的样本，用于网络模型的训练以及测试。

本实验中的数据采集部分的工作是在 Hadoop 分布式集群上完成的。而神经网络的搭建与训练是基于 TensorFlow 实现的。TensorFlow 是谷歌于 2015 年 11 月新开源的深度学习平台。TensorFlow 提供了非常丰富的深度学习相关的 API 接口，兼容性好，并且 TensorFlow 与 Numpy 完美结合，这使大多数精通 Python 数据科学家很容易上手。它能够进行自动求导，大大节约了编程的时间开销。并且支持 GPU，大幅度提高神经网络的训练效率。TensorFlow 的具体介绍本文不再赘述。

3.6.2 网络实现

在数据集划分上，为了防止过拟合，本文采用了十字交叉验证的方法对网络进行训练和测试，以确保实验结果的可靠性。并在数据输入网络之前，对数据进行归一化处理，以防止出现梯度爆炸或者梯度弥散。在本实验中采用 mini-batch SGD 对网络的目标函数进行优化，batch 的大小为 1024。经过多次实验发现当迭代次数到 10 次左右时，网络基本达到收敛。

根据 Montavon[52]提出的理论，本实验在对神经网络的权重矩阵进行初始化时，采用随机高斯分布的初始化，均值为 0，方差为 0.01。矩阵中元素的取值范围为

$-\sqrt{6/(fanin+fanout)}$到$\sqrt{6/(fanin+fanout)}$，其中$fanin$和$fanout$分别表示相邻两个网络层中的神经元的个数，如图 3-2 所示 fanin 表示的是输入层中的神经元个数，fanout 则表示的是 h_1 中神经元的个数，以此类推。

为了在实验中防止过拟合情况的发生，在实验中采用了 Hiton 在 2012 年提出来的 Dropout 方法[55]，该方法简单来说就是在深度神经网络的训练过程中，对于神经网络单元，按照一定的概率使其暂时停止工作。也就是说在神经网络的训练过程中，网络每次都只考虑一部分特征对最后结果的影响，从而减小过拟合情况的发生，这与随机森林的思路有些相似。本文使用该方法的思路在网络模型的隐含层中使用的 Dropout 机制，其中停止工作的概率值设为 0.5。

3.6.3 实验结果与分析

在深度语义模型中，有几个关键的参数对系统的性能影响较大。其中包括，平滑因子 γ 、输出层输出特征维度、神经网络层数以及学习率 λ 。对于平滑因子针对这几个参数，本文分别设置了不同数值进行多组对比实验，并分别从 RMSE、AUC，以及准确率和召回率等几个评价指标上对他们进行了有效性评价和分析，最后结果如下所示。

图 3-3 给出了不同平滑因子 γ 对系统性能的影响，从图中可以看出，当平滑因子 γ 的取值从 1 逐渐递增到 30 时，基于深度语义模型推荐算法的 AUC 指标是先递增后减小的，而 RMSE 指标是先减小后增大的。当 γ 取值在 15 附近时，AUC 和 RMSE 分别取得了最大值和最小值，这个时候整个推荐系统的性能最佳。

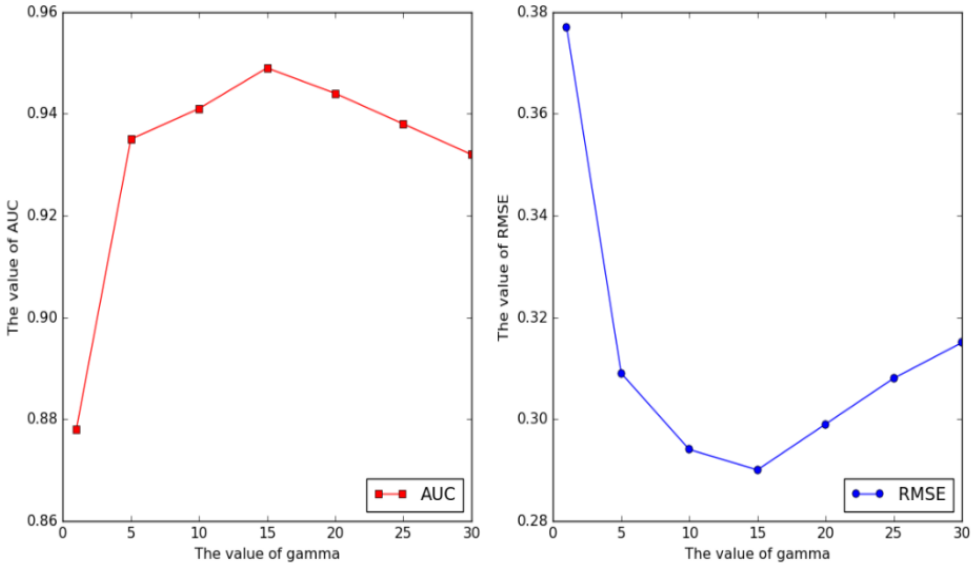

图 3-3 不同的平滑因子 γ 取值对系统性能的影响

表 3-3 不同的输出特征维度对系统性能的影响

输出特征维度	AUC	mAP	RMSE
60	0.929	0.938	0.33
120	0.942	0.948	0.307
180	0.943	0.948	0.302
240	0.944	0.949	0.299
300	0.945	0.950	0.297

输出特征维度反映的是用户与视频之间潜在的特征关系的数量，从表 3-4 可以看出，当随着输出的特征维度从 60 到 300 逐渐递增时，本算法的 AUC 指标和 mAP 指标在逐渐增加，而 RMSE 指标在逐渐减小，这说明系统的性能是随着输出特征的维度增加而提高的。但是，当输出特征在大于 180 之后，系统的性能提升有限，反而对神经网络训练的时间开销较大，因此综合考虑输出特征维度在 240 的时候最佳。

表 3-4 不同的神经网络层数对系统性能的影响

神经网络层数	AUC	mAP	RMSE
3	0.908	0.907	0.357
4	0.916	0.929	0.345
5	0.944	0.949	0.299

从表 3-4 中可以看出，随着神经网络层数的递增，系统的的评价指标 AUC 值和 mAP 值在逐渐变大，RMSE 在逐渐减小，这说明神经网络层数约深，本算法的性能越好。这是因为在数据量充分的情况下，神经网络的层数越深它对视频以及用户的特征提取越充分，能够从中获取到较为关键的特征，从而能使系统的性能得以提升。

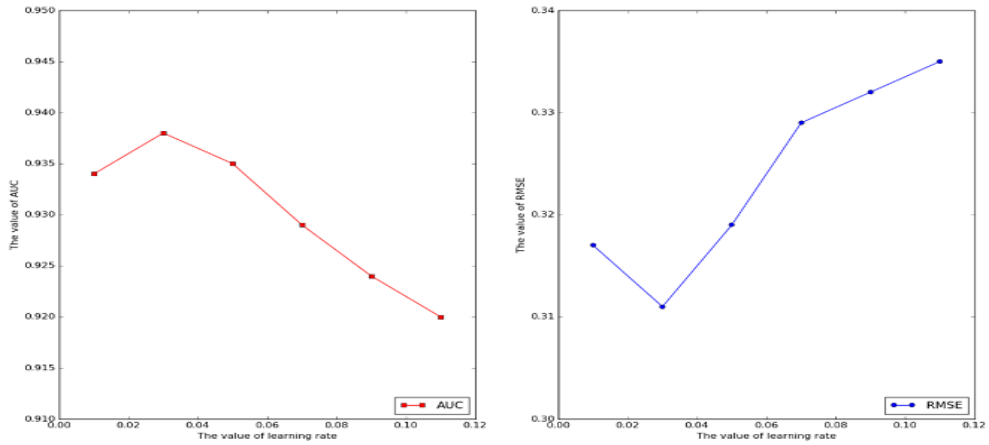

图 3-4 学习率对系统性能的影响

学习率决定了梯度下降的幅度，学习率过小使得神经网络的训练速度较慢，而学习率过大则会使得训练结果达不到最优，引起收敛过程中的反复震荡。本文使学习速率从0.01 以步长 0.02 的速度增长到 0.11，得到了 6 组 AUC 和 RMSE 的结果并分别画出了它们变化的曲线图，如图 3-4 所示。从图中我们可以看到，随着学习率的递增，本算法的 AUC 指标是先增大后减小，而 RMSE 指标是先减小后增大，并且当学习率在 0.03 附近时，AUC 和 RMSE 同时达到最大值和最小值。这说明当学习率取得 0.03 时，系统此时的性能最佳。

除此之外，为了更加客观地展示本算法的有效性，接下来本文将本算法与基于 random 的推荐，基于 TF-IDF 的推荐算法，基于 UserCF 的推荐算法以及基于 ItemCF 的推荐算法进行了对比，结果如下所示。

其中，图 3-5 给出了四个算法的 ROC 曲线示意图，从曲线形式可以看到，基于深度语义模型的推荐算法的 ROC 曲线完全包住了其他三个算法的 ROC 曲线，这说明它在预测用户是否会观看目标视频的精准度上要明显地高于 itemCF、userCF 和 TF-IDF 三个算法，具有较大的优势。

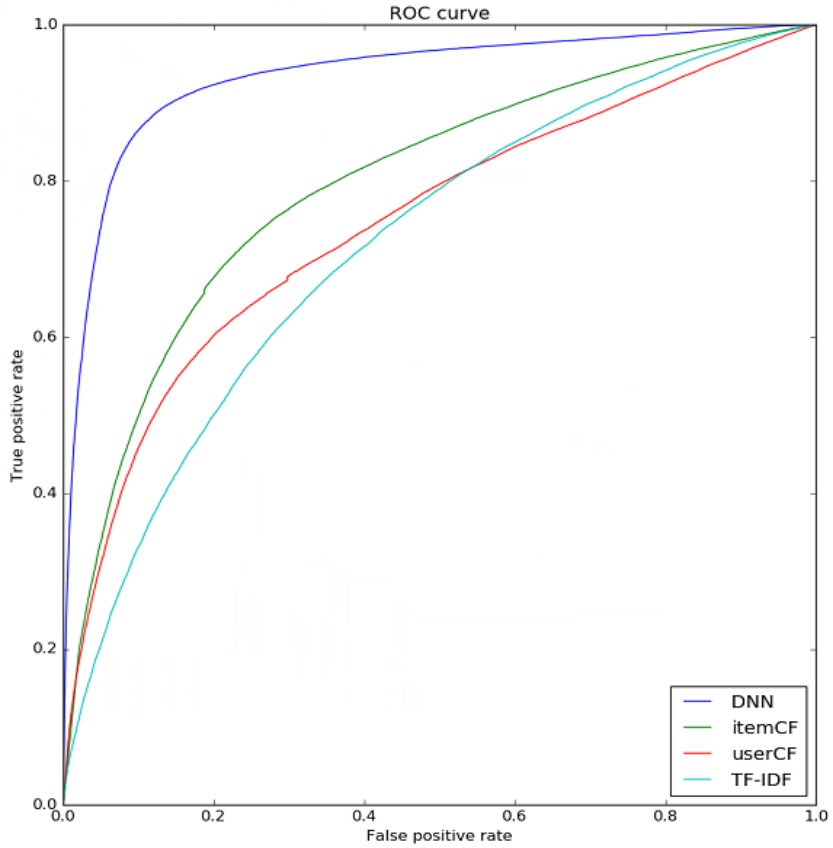

图 3-5 不同算法的 ROC 曲线示意图

表 3-5 不同算法之间的性能指标对比

推荐算法	AUC	mAP	RMSE
DNN	0.935	0.937	0.309
ItemCF	0.801	0.803	0.450
UserCF	0.750	0.764	0.486
TF-IDF	0.720	0.713	0.474

从表 3-5 可以看到，本文所提出的基于深度语义模型的推荐算法的 AUC 指标和 mAP 指标都要高于其他的三个算法，而 RMSE 指标则要低于它们。这说明本算法的性能要优于对比的三个算法，具有很大的可行性。这主要因为本算法在基于文本内容的基础上，通过深度神经网络充分地提取用户和视频的内容特征，深度地挖掘了视频与用户之间的语义相关性，从而取得了非常不错的推荐效果。

3.7 本章小结

本章主要是针对协同过滤算法中冷启动问题和数据稀疏性问题，提出了一种基于内容的推荐算法，该算法结合深度神经网络深度挖掘用户和视频的内在联系，以此作为依据进行推荐工作。该算法首先利用 TF-IDF 算法对视频中的文本描述信息进行了特征抽取，提取文本内容中较为重要的关键词作为视频的标签信息，作为视频的特征描述。然后根据用户对推送视频的历史行为记录，分析用户的偏好，利用视频向量计算得到用户的特征向量，进行用户建模。接着搭建深度神经网络模型，将计算得到的视频和用户向量作为神经网络的输入，通过用户的历史点击行为记录，进行有监督的网络学习。最后根据训练好的深度模型框架进行用户的对视频的评分预测，生成推荐结果并反馈给用户。

本章最后，进行了实验设计，并对结果进行了对比分析。结果表明，本章所提出的基于深度语义模型的推荐算法具有很好的推荐效果，从各项指标上来看，和同比的 TF-IDF 算法、基于 userCF 以及基于 itemCF 的推荐算法要有明显的优势。

第 4 章　基于概率语言模型的视频个性化推荐系统

基于 TF-IDF 算法得到的视频向量以及用户向量存在有一定的局限性：一方面，是因为它只是单纯的通过词频统计信息去评估某个关键词在文中重要性，只是将词语看成了独立的个体，并没有从语义层面去分析词语之间的关联信息。例如猫和狗两个词，用 TF-IDF 建立的特征向量，是无法知晓猫和狗的关联信息（比如：都是动物，都有四条腿等），这会使的 TF-IDF 算法对视频的内容特征表示不充分。另一方面，用 TF-IDF 算法构建的用户向量和视频向量，其数据非常稀疏。因为在实际应用中，往往一个的标签池是比较大的，而对单一的某个用户和视频而言，所用的标签相对而言就会显得很少，尤其是对视频来说，它的标签很多情况下只有十来个左右，因此在基于 TF-IDF 算法得到的用户和视频的特征向量是非常稀疏的，过于稀疏的输入对神经网络训练是非常不利的，可能会使神经网络很难学到有用的东西。

基于以上原因，所以希望能找到一种的视频和用户向量的表示方法，既能描述用户与视频之间的内在联系，又能使得到的特征向量是低维非稀疏的。受到了词向量模型[53]的启发本章提出了一种基于概率语言模型的用户和视频特征建模算法，并将该算法用于视频的个性化推荐。

4.1 词向量模型

在自然语言处理领域，词语的分布式表示即词语的向量化[53]是一个非常热门的研究课题。这是因为词向量能够以一种量化的形式去描述词语与词语之间的相关信息，它有利于自然语言处理问题数学建模，并且方便机器对自然语言进行理解和学习。

近些年来，提出了很多关于词语分布式表示的算法，其中最著名的便是 Mikolov 等人在 2013 年提出来的 word2vec 算法。该算法因其快速，高效而闻名。它通过构建神经网络对语料库中的词语进行学习和训练，利用上下文环境来分析词语之间的内在联系，并最后以一种分布式表示的形式对词语的语义特征进行描述。

Word2vec 模型包含两种框架，分别为连续词袋模型（CBOW）和 Skip-Gram 模型。其结构如图 4-1 所示。

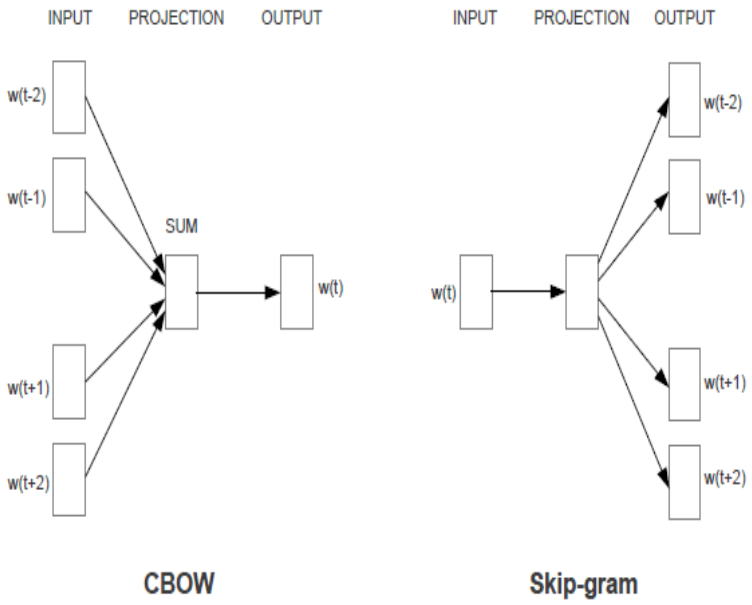

图 4-1 word2vec 的两种模型框架图

从算法上来看，这两种方法非常相似。其区别在于 CBOW 模型是根据上下文来预测目标词汇，而 Skip-Gram 模型是根据目标词汇来预测上下文。Skip-Gram 模型采取 CBOW 的逆过程的动机在于：CBOW 算法对于很多分布式信息进行了平滑处理（例如将一整段上下文信息视为一个单一观察量）。很多情况下，对于小型的数据集，这一处理是有帮助的。相较之下，Skip-Gram 模型将每个“上下文-目标词汇”的组合视为一个新观察量，这种做法在大型数据集中会更为有效。

为了比较直观地展示词向量分布式表示效果。本文利用 word2vec 算法对 Text8 语料进行训练，并对结果进行检测，在图 4-2 中给出了一些有意思的查询结果图。该结果是按词语向量之间的余弦相似度的大小进行排序，越靠前的词语说明训练结果得到的语义越相近。从结果显示来看，基于 word2vec 的词语向量化表示在相当程度上能够反映出词与词之间的内在联系。

本文受到了 word2vec 算法的启发，在 word2vec 的基础上进行了改进，加入了用户信息，提出了一种用户以及视频向量训练的网络模型，通过该神经网络的训练对用户以及视频进行一种分布式表示，挖掘用户与视频之间的潜在联系，并在此基础上进行后续的个性化推荐工作。

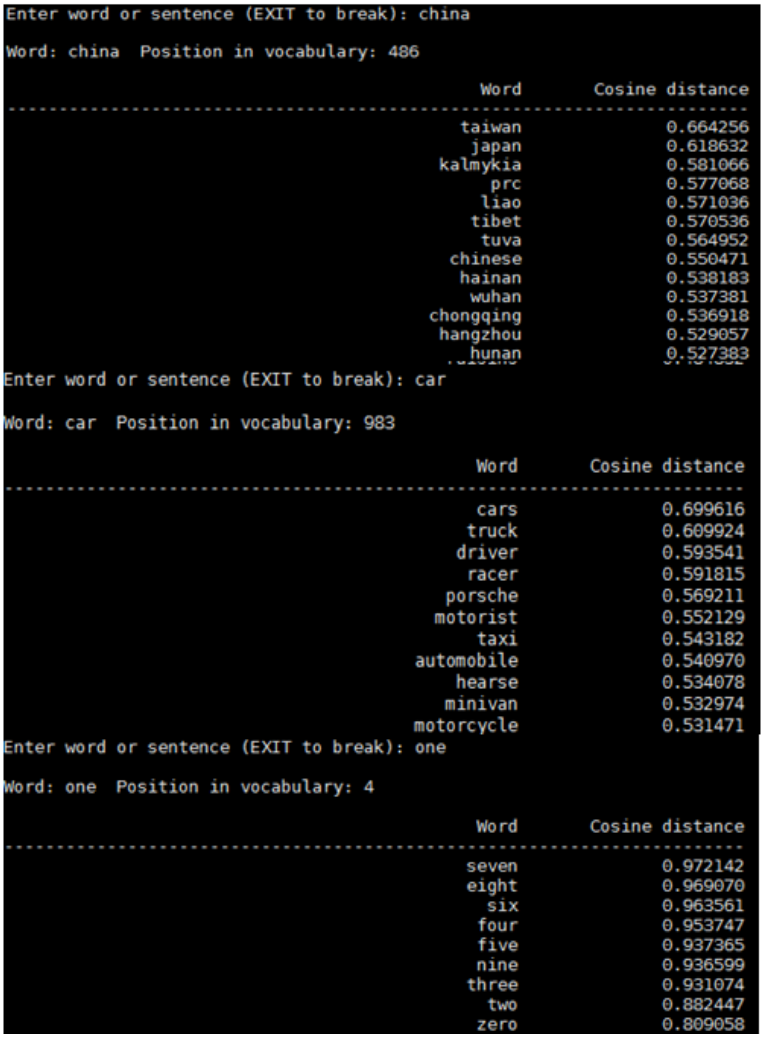

图 4-2 相似度查询结果示意图

4.2 基于概率语言模型的算法原理及实现

4.2.1 算法原理

从 4.2 节可以看出，word2vec 算法可以获取词语在语义层面的特征，能够挖掘出词语与词语之间的内在联系。那么是否可以也可以利用该算法分析出用户与视频之间的内在联系呢？

在回答之前，我们首先思考一个问题：什么是自然语言？语言其实在广义上可以理解成是一种特殊符号，因此自然语言其实就是一组符合一定自然规律的符号序列。在这里，所谓的符合自然规律是指用户说的话、写的字等符合一定语法结构，逻辑规则。我们将此类比到用户观看视频的行为，如果除开观看视频之间的时间间隔，把用户观看视频看成是一个连续的过程，那么用户所观看的视频的序列实际上也是符合某种规则和逻辑的，它反映出了视频与视频之间的某种联系以及用户的某种行为习惯。在实际中生活

中，用户连续观看的视频之间的很多时候都是具有很强的相关性的，比如说连续追动漫剧集，连续看某个明星的相关视频等。因此，用户观看视频这一行为在某种程度上和人说话、人写字的这类行为是等价的。如果，给不同的视频以不同 id 作为它的符号，那么一组用户所观看的视频的 id 序列的其实和自然语言一样是一组有序的符号序列。因此，可以将用在自然语言处理上的某些方法应用到视频序列上来。

但是，用户观看的视频序列与自然语言相比是有很大的不同的。主要是因为不同的人的兴趣爱好会有很大的区别，适用于某个用户的逻辑规则在另外一个用户上或许就行不通了，但是自然语言不同，因为语言本身拥有一套比较统一的语法规则，在不同的人身上它的变化不是很大。基于视频序列与自然语言上的不同的原因，本文在 Mikolov 提出的 CBOW 模型基础上加以改进，在网络中加入了用户的信息，同时对视频和用户向量进行训练，使得训练得到向量即包含有视频之间的关联信息，同也包含有用户与视频之间的关联信息。下文将详细介绍该算法的实现过程。

4.2.2 算法实现

对于某个特定用户 u，给出一组视频序列 $\{v_1, \dots, v_n\}$，和自然语言一样想要判断这组视频序列是否是该用户想要观看的视频序列，我们可以使用它们的联合概率 $\mathrm{P}(v_1, \dots, v_n|u)$ 表示[54]，当它超过某个阈值时则认为它是，反之则不是。从物理意义上来看，可以理解成是用户连续观看这组视频的可能性大小。通过贝叶斯公式将该式子展看可以得到：

$$\mathrm{P}(v_1, \dots, v_n|u) = \mathrm{P}(v_1|u)\mathrm{P}(v_2|v_1, u) \cdots \mathrm{P}(v_n|v_1 v_2 \dots v_{n-1}, u) \tag{4-1}$$

如果令 $\mathrm{Click}_i = (v_1, \dots, v_{i-1})$，上则可以表示成：

$$\mathrm{P}(v_1, \dots, v_n|u) = \prod_{i=1}^{n} \mathrm{P}(v_i|\mathrm{Click}_{i-1}, u) \tag{4-2}$$

于是这个联合概率就表示成了一组条件概率连乘的形式。如果我们把Click_i 理解成是用户 u 在观看视频 v_i 前的历史观看行为记录。因此 4-2 式实际上是利用 u 的历史行为来分析判断他看下一个视频 v_i 的概率，这与我们推荐思想是相符的。

因此，就可以将推荐问题转换成对条件概率 $\mathrm{P}(v_i|\mathrm{Click}_{i-1}, \mathrm{u})$ 的计算问题。通过上面的定义，我们知道它反映了用户 u 观看视频 v_i 的概率是受到过去过去所有观看行为影响的。然而在实际中，用户的兴趣爱好其实是在不断变化的，用户当前观看行为其实只与最近一段时间的观看行为相关。因此我们可以将上述的$\mathrm{Click}_i = (v_1, \dots, v_{i-1})$变成 $\mathrm{Click}_i = (v_{c-\mathrm{m}}, \dots v_{c-1} v_{c+1} \dots, v_{c-m})$, 其中$m$视为一个窗口值，表示视频$v_c$只受最近观看

的 $2m$ 部视频的影响。

然后令Click_i 可以用向量 v_{Click}表示，定义条件概率如下：

$$\text{P}(v_i|\text{Click}_i, u) = \frac{\exp\,(v_i \cdot (v_{\text{Click}}, u))}{\sum_{k=1}^{|V|} \exp\,(v_k \cdot (v_{\text{Click}}, u)} \tag{4-3}$$

此条件概率代表的就是根据用户 u 的历史观看记录判断用户观看视频 v_i 的概率的大小。

构建神经网络，其框架结构如图 4-4 所示。

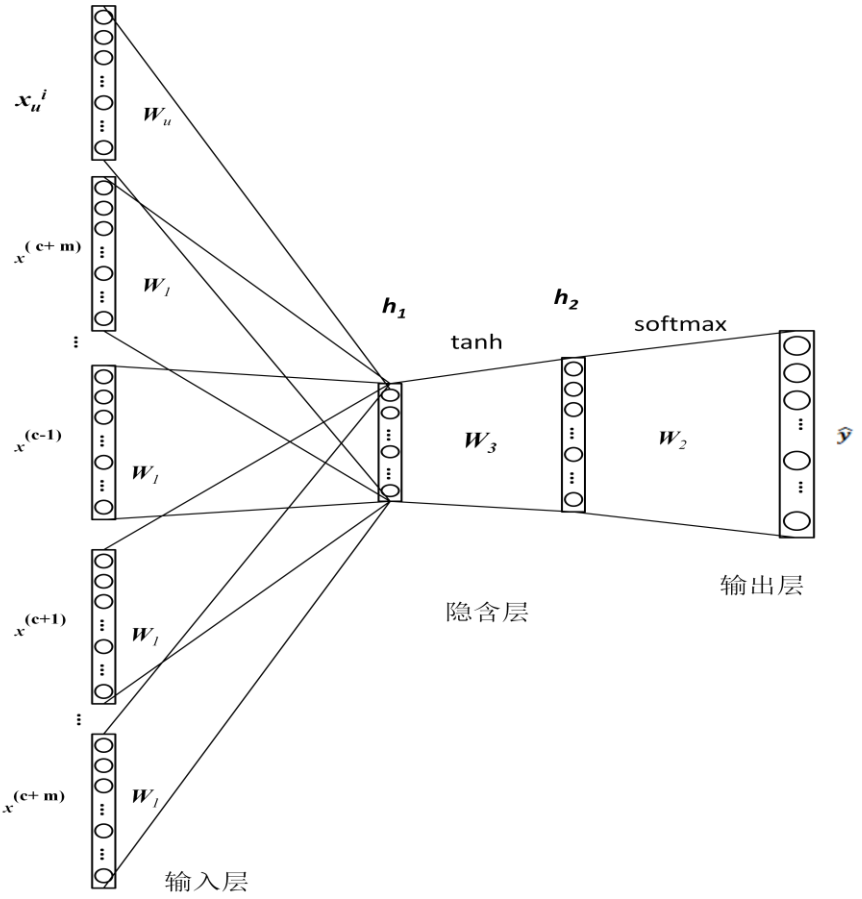

图 4-3 算法网络模型框架

定义符号如下： $x^{(c)}$ 以及 u 为输入向量，是一组one-hot编码的向量，维度大小为分别为 $|V|$和$|U|$，$|V|$表示视频集大小，$|U|$表示用户集大小；$\text{W}_1 \in \mathbb{R}^{\text{n}\times|\text{V}|}$ 和 $\text{W}_2 \in \mathbb{R}^{\text{n}\times|\text{V}|}$为视频向量矩阵，其中 n 表示定义的视频向量的维度，W_1是输入视频向量矩阵，它的每一列表示一个视频向量v_i，W_1是输出视频向量矩阵，它的每一行表示一个视频向量v_i。$W_u \in \mathbb{R}^{\text{n}\times|\text{U}|}$是用户向量矩阵，每一行表示一个用户向量 u_i。W_1，W_2，W_u是我们需要学习的参数。

该网络训练的流程为：

1）对用户 u_i 以及他所观看观看过的视频$(v_{c-\mathrm{m}},\dots v_{c-1}v_{c+1}\dots,v_{c-m})$进行 one-hot 编码，即向量 u_i 只在第 i 个位置为 1，其余位置均为零，视频向量亦如此。

2）计算用户 u_i 的视频向量以及每个历史点击视频的视频向量。

$$u_i = \mathrm{W}_u x_u^i \tag{4-4}$$

$$(v_{c-\mathrm{m}} = \mathrm{W}_1 x^{(c-m)},\dots v_{c-1} = \mathrm{W}_1 x^{(c-1)} v_{c+1} = \mathrm{W}_1 x^{(c+1)}\dots,v_{c-m} = \mathrm{W}_1 x^{(c+m)}) \tag{4-5}$$

3）将上述点击视频向量的算术平均作为整个历史行为向量即：

$$v_{\mathrm{Click}} = \frac{v_{c-\mathrm{m}}+\cdots+v_{c-1}+v_{c-1}+\cdots+v_{c+\mathrm{m}}}{2m} \tag{4-6}$$

4）再将 v_{Click} 与用户向量 u_i 进行拼接，得到隐藏层 h_1 中的向量：

$$h_1 = (u_i, v_{\mathrm{Click}}) \tag{4-7}$$

5）将 h_1 进行一次非线性变换，得到 h_2：

$$h_2 = tanh\,(W_3 h_1 + b_1) \tag{4-8}$$

其中$\tanh(\mathrm{x}) = \frac{1-\mathrm{e}^{-2\mathrm{x}}}{1+\mathrm{e}^{-2\mathrm{x}}}$。

6）再得到一组对各视频的评分向量：

$$z = W_2 h_2 + b_2 \tag{4-9}$$

7)通过 softmax 回归将上面的评分向量转换成概率分布 $\hat{\mathrm{y}}$，并作为预测输出，其中：

$$\hat{y}_i = \mathrm{p}(v_i|\mathrm{Click}_i, u) = \frac{\exp(z_i)}{\sum_{k=1}^{|V|}\exp(z_k)} \tag{4-10}$$

我们希望得到概率分布 $\hat{\mathrm{y}}$ 能够非常地接近真实值 $y = x^{(c)}$，因此我们使用交叉熵作为目标函数，定义如下：

$$H(y,\hat{y}) = -\sum_{i=1}^{|V|} y_i log\hat{y}_i \tag{4-11}$$

因为 y 是一个 one-hot 向量，它只有在 i 处的值为 1，其他地方值为 0，所以上式可以简化为：

$$H(y,\hat{y}) = -log\,\hat{y}_i \tag{4-12}$$

因此我的优化目标就是最小化上面的交叉熵：

$$\mathrm{minJ} = -log\,\hat{y}_i = -z_i + \log\sum_{k=1}^{|V|}\exp(z_k) \tag{4-13}$$

从上式可以看到，由于每更新一次视频向量就需要对整个视频集遍历一次，因此是非常耗时的，需要对上面的目标函数进行变化。这里采用负采样的策略，重新建立目标函数如下：

$$\text{minJ} = -log\,\sigma(w_{c-m+j}^{2} \cdot h_2) + \sum_{k=1}^{K} log\,\sigma(-\widetilde{w}_k^2 \cdot h_2) \qquad (4\text{-}14)$$

其中，w_{c-m+j}^{2}表示视频向量矩阵 W_2 中的视频向量 w_{c-m+j}，$\widetilde{w}_k$是从通过负采样得到的 k 个负样本。此处目标函数的重建可以参照 word2vec 中 CBOW 模型的目标函数构建。最后利用 mini-batch SGD 优化算法对上面的目标函数进行优化，最后经过训练得到用户的向量矩阵 W_u 以及视频的向量矩阵 W_1，并基于得到的用户和视频向量进行视频的个性化推荐。

4.3 基于概率语言模型的推荐流程

4.3.1 数据获取以及数据预处理

首先，根据需求通过 Hadoop 大数据平台获取所需要的相关数据，包括两方面的内容：第一个是用户的历史观看行为记录，用于视频和用户建模；第二个是用户对推送视频的历史点击行为的数据，用于对模型进行测试。

数据预处理工作主要是对数据进行清洗，检查数据一致性，过滤掉信息缺失数据和无效数据。然后整理数据，获得每个用户的 ID 信息以及对应用户当天所观看的视频 ID 信息，并统计用户以及视频信息，生成用户索引以及视频索引。

4.3.2 视频及用户向量建模

用户准备好的视频及用户 ID 信息，根据 4.3 节中所提出的算法，搭建并训练神经网络模型，提取视频与用户之间以及视频与视频之间的相关性特征，得到用户和视频的分布式特征表示，以此作为视频及用户的向量模型，用于描述视频及用户的特征信息。

基于概率语言模型算法所得到的用户及视频向量，它们是低维且非稀疏的，并且它们在一定程度上能够反映出视频之间以及用户与视频之间的内在联系，所以能够以此作为依据进行后面视频的推荐工作。

4.3.3 生成推荐列表

当利用上述网络训练得到用户以及视频的向量表示后，便可以通过计算用户向量与视频向量之间的余弦相似度的值，来表示用户对目标视频的兴趣度的大小，然后将得到的结果按从大到小的顺序进行排列，取值靠前的 Top-N 个视频进行推荐，并将结果反馈

给用户。

以上便是基于概率语言模型的视频个性化推荐算法的基本流程介绍，其流程框架可以见图 4-4 所示。

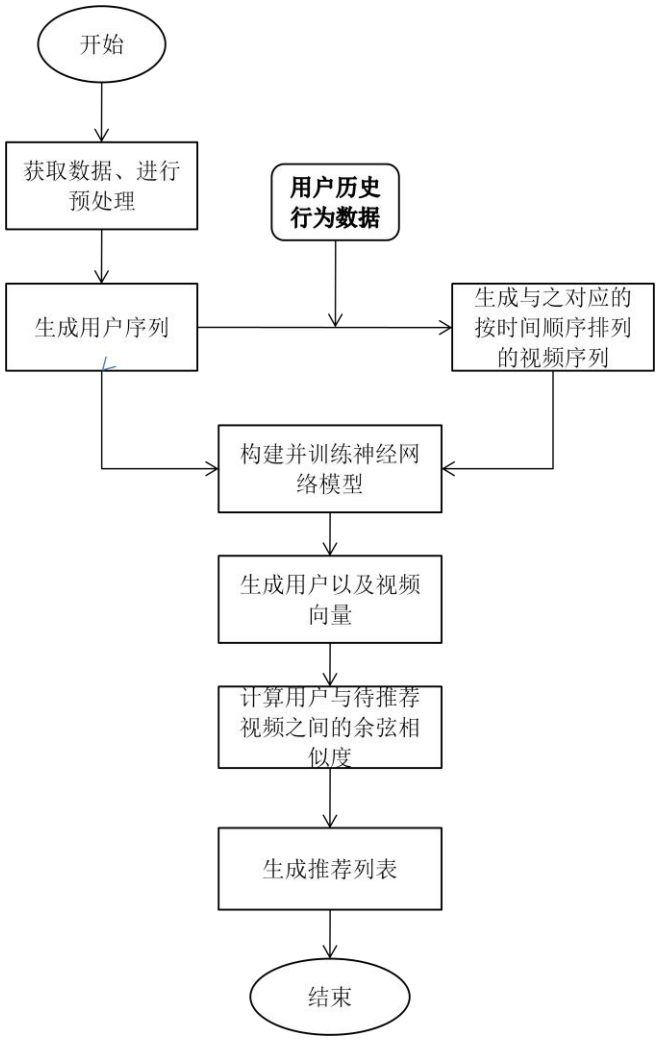

图 4-4 基于概率语言模型的视频推荐流程示意图

4.4 实验设计与结果分析

4.4.1 实验数据和实验环境

在本算法中，所使用的实验数据主要有：1）用户过去 15 天所有的历史观看行为数据，包括 3 千多万个用户每天的观看历史记录，相关示例如表 4-1 所示，其中用户 ID 是经过脱敏处理后的唯一标识，视频信息的格式为“来源：视频 ID：播放时长”；2）用户的对推送视频的点击行为记录数据，这部分数据与第三章中所使用的数据内容相同，主要是用于实验结果测试。

表 4-1 用户历史观看行为数据示例

用户 ID	视频信息	时间
User132144	HomeRec:a0306tiqf7l:0.70， HomeRec:q03064fhljn:0.23	2016-06-01 09:31:22
User331943	unknow:j0020dwz3bp:1.00， unknow:k0020pow63n:1.00， HomeRec:p0020dktii1:0.81	2016-06-02 21:58:02
User562311	unknow:v0307kfw7ci:1.00， unknow:w0307f1kp5e:0.93	2016-06-01 09:31:22
User948913	Push:206709:1， UserCenter:n0305jq2d80:0.97	2016-06-03 12:41:25

本实验的所使用的实验环境包括：1）Hadoop 分布式大数据平台，用于收集和获取用户日志信息。2）TensorFlow 深度学习平台，用户神经网络的搭建和训练工作。

4.4.2 实验结果与分析

通过上文的介绍可知，本推荐算法的核心在于基于改进的 CBOW 模型进行视频向量以及用户向量的建模部分，向量训练结果的好坏直接关系到后续推荐系统的性能。为了对该算法的有效性进行展示。表 4-5 给出了基于改进的 CBOW 模型得到的视频向量的 Top-5 的查询结果图，越靠前的表示越相似。

表 4-1 目标视频的最近邻集查询结果图

目标视频 ID	目标视频内容	最相似的前 5 个视频 ID	最相似的前 5 个视频内容
n0301234bo6	杨幂刘恺威再被传离婚 刘丹：以后不回应 烦死了	y0300g8c9z3； t03013azpmj； s0020ajulf3； q03016hhed1； e0019l9qzjq；	杨幂刘恺威再被曝离婚 这次是金融界透出的消息； 现场：杨幂发声回应离婚传言 与黄轩拍亲密戏不尴尬； 被传代孕、离婚、不雅视频 为啥杨幂总招黑； 杨幂亮相未戴婚戒 回应离婚传闻称现在状态很好； 杨幂离婚鹿晗结婚生子 2015 娱乐圈谣言都在这；
w0020g1vmxm	前方直击骑士训练 乐福苦练三分詹皇偷懒观战	o0020ywovs1； u00205s6xri； e0020ds1mug； r0020hfyg6w； e0020rmim47；	前方直击骑士训练 欧文杂耍运球飙干拔三分； 前方直击骑士备战总决赛 莫兹戈夫防守敏捷送大帽； 前方直击骑士备战总决赛 欧文单打训练师干拔精准； 詹姆斯与欧文训练比拼三分 欧文手感颇佳完爆詹皇； 前方直击：勇士老板再度鼓励汤普森 你会再次接管比赛；
h0019px66y0	迷你特攻队_08	g0019kt3ch7； h0019j5v5eh； t00192ktzk9； k0019t71uf2； j0019ppaiyr；	迷你特攻队_09； 迷你特攻队_07； 迷你特攻队_06； 迷你特攻队_10； 迷你特攻队_05；

从上表可以很直观的看到，无论是内容格式非常规范的电视连续剧，还是用户自定义上传的 UGC 类视频，通过对目标视频的最近邻查询的其结果在内容都是非常相似的。

这说明了本文提出的改进的 CBOW 算法所得到的视频向量能够在一定程度上反映出视频之间的相关性，通过该算法生成的视频向量能很好地对视频特征信息进行描述。

如此之外，本文将本算法中所得到的视频向量进行了聚类，并将最后得到的结果降维到二维平面，其结果如图 4-5 所示，图中相同颜色的坐标点代表的是相同频道的视频。从图中可以很直观的看出，基于改进的 CBOW 算法所得到的视频向量具有很理想的聚类效果，相同频道的视频相对来说分布都比较集中，而不同频道之间的视频则相隔较远。这也从另外一个角度说明了通过该算法的训练得到的视频向量体现了视频与视频之间的相关性，越相似的视频则它们的距离越接近，反之则越远。

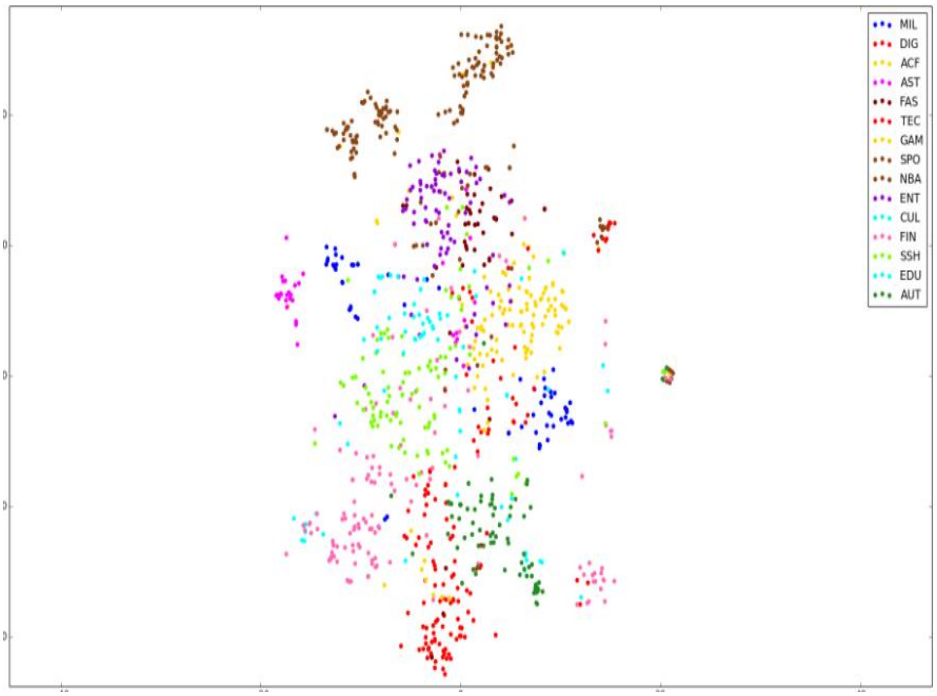

图 4-5 视频向量二维平面效果图

接下来，为了更加客观地展示基于概率语言模型推荐算法的有效性，本文将该算法与传统的 TF-IDF 算法、UserCF 算法以及 ItemCF 算法进行了对比。其结果如下所示。

表格 4-2 各算法的性能指标对比

算法名称	AUC	mAP	RMSE
基于概率语言模型	0.830	0.830	0.425
ItemCF	0.751	0.768	0.485
UserCF	0.769	0.779	0.465
TF-IDF	0.670	0.704	0.535

表 4-2 给出了四种不同算法的 AUC、mAP 和 RMSE 指标的结果。从表中可以看到，TF-IDF 算法的各项性能指标在四种推荐算法中是最差的，这主要是因为 TF-IDF 算法的局限性所致的，因为它只考虑了词频的统计信息，词语之间的关联信息则被它所忽略，从而使得用户及视频的特征表示不精准，影响推荐效果。而 UserCF 以及 ItemCF 的性能指标比较接近，相差不大，但都要低于本算法，这是因为协同过滤的推荐算法受到比较严重数据稀疏性问题的影响，尤其是在数据量很大的情况下，这种影响更加明显，使得它的性能受到较大的制约。而本算法的三种指标都要优于其他算法，这说明本算法在推荐准确度上是要比另外三种算法好。这是因为本算法所提出的视频以及用户向量建模算法，充分考虑了视频之间以及视频与用户之间的相关性，利用深度学习对视频以及用户的相关特征进行深度挖掘，同时在网络训练过程中巧妙地借鉴 CBOW 模型的某些特点，直接对用户以及视频进行向量训练，避免了传统协同过滤需要构建大规模的评分矩阵所带来的数据稀疏问题的困扰。因此本算法具有很高的可行性。

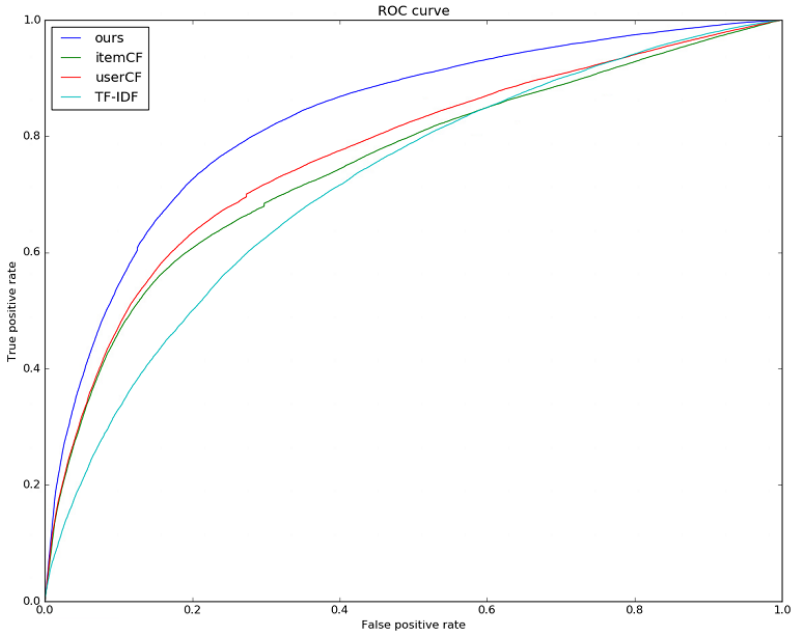

图 4-6 各算法的 roc 曲线图

图 4-6 给出了四种算法的 ROC 曲线图，从图中可以看到，本算法的 ROC 曲线要优于其它三种算法，这说明了本算法对用户点击行为能够有很不错的预测结果。这是因为本算法在对用户和视频向量的进行建模的过程中充分地利用他们之间的关联信息，因而训练所得到的用户向量与视频向量能够比较理想地反映出他们之间的相关性大小。这使得本算法能够比较准确地预测用户对目标视频的喜爱程度。这从另一方面说明了本算法具有很高的可行性。

4.5 本章小结

本章针对 TF-IDF 算法的局限性，提出了一种基于概率语言模型的用户以及视频向量训练算法。该算法与 TF-IDF 算法相比，它所提取的用户向量以及视频向量能够很好地反映出视频与视频之间以及用户与视频之间的内在联系。在此基础上，基于该算法所得到的视频与用户向量进行视频推荐。最后通过实验结果对比情况来看，本算法的推荐效果比要比传统的 IF-IDF 算法、UserCF 算法和 ItemCF 有着明显的优势。

第 5 章　总结与展望

5.1 论文工作总结

随着互联网和视频多媒体技术的快速发展，人们的生活娱乐方式也在发生着改变，观看网络视频成为了人们业余生活消遣的一个重要途径。但随着视频量的和用户量的激增，网络视频播放领域也必然要面临信息过载问题。本文主要是针对视频业务领域对的个性化推荐系统展开了研究。

近些年来，深度学习成为了一个最热门的研究方法，它在语音识别、图像处理和自然语言处理等领域取得了巨大成就。但是基于深度学习的个性化推荐系统的研究工作还处于探索阶段。在此基础上，本文利用深度学习在自然语言处理上取得的成就，结合传统的基于内容和基于协同过滤的推荐算法对推荐系统进行了深入的探讨。提出了两种基于深度学习的视频个性化推荐算法。本文的主要工作内容包括以下几个方面：

1）论文首先介绍了推荐系统的相关背景并对其研究意义进行了阐述，然后简要地介绍了目前国内外的研究现状。紧接着从相关技术背景出发，对推荐系统的整体框架以及当前主流的推荐算法进行了详细介绍，并总结了各自算法的优缺点。最后简要概述了推荐系统的性能指标。

2）针对协同过滤推荐算法的中的冷启动问题和数据稀疏性问题，提出了一种基于深度语义模型的个性化推荐算法。该算法结合了深度学习在基于内容的推荐算法的基础上进视频推荐设计。具体来说，首先对视频的文本描述信息利用 TF-IDF 算法进行关键词提取，并将较为重要的关键词作为视频的标签；然后构建视频向量，以标签的 TF-IDF 值作为向量中元素的值，以此来描述视频的特征信息；接着根据用户的历史点击行为记录，计算所观看的视频的算术平均作为用户的向量表示，用次来描述用户的特征信息；然后搭建深度神经网络，利用用户的点击行为数据作为训练样本，对模型进行有监督地训练。最后拿训练好的模型进行视频评分预测，以此作为视频推荐的依据。最后的对比实验结果显示，本算法较 TF-IDF、UserCF 和 ItemCF 的推荐效果要更佳。

3）提出了一种基于概率语言模型的视频个性化推荐算法。该算法从用户的观看行为出发，将用户所观看的视频序列类比成一组有序自然语言。并将每一个视频项看做是

语言中的一个词语。在此基础上，对 word2vec 算法中的 CBOW 模型进行改进，引入用户信息视频及用户本身直接向量化表示，不需要通过分析视频的内容信息进行构建。当获得视频和用户向量后，根据用户与视频向量之间的余弦相似度进行视频的推荐。该算法通过巧妙的类比，直接将视频序列理解成自然语言。并且对 word2vec 算法进行改进，引入了用户信息，使得所得到的视频向量既体现了视频之间的相似性，同时又反映了与用户之间的相似性。本算法较之传统的协同过滤算法不需要构建庞大的评分共现矩阵，避免了协同过滤中的数据稀疏性问题。并且在实验结果上显示，该算法的性能较之于 TF-IDF 算法、UserCF 算法和 ItemCF 的算法要好。

5.2 未来展望

个性化推荐系统是一个非常有意思的研究领域，它与人们平常生活息息相关，具有非常大现实意义。随着深度学习在各个领域中的迅速突破，我相信在未来基于深度学习的个性化推荐系统必然有着广阔的前景。本文在这里只是结合深度学习做了一些小的尝试，其工作还有大量的不足。未来可深入的地方有：

1）本文的推荐工作都是基于文本内容，而视频并非只有文本信息，它还包括了语音和图像等。我们可以利用深度学习在语音识别和图像处理上的相关技术，对视频的图像和语音信息进行特征提取，以便获取更加完善的视频特征信息。进而有助于发现更多的规律。

2）本文对用户的建模是基于视频内容基础之上的，我们还可以考虑加入用户本身的基本特征信息。如性别，年龄，地域等，这样有助于丰富和完善用户的特征描述。使得推荐更加精准。

3）本文的推荐设计中没有考虑视频的分类信息，现在视频网站都会对视频进行分类，如何体育频道、娱乐频道、新闻频道等，这些信息也是视频非常重要的特征。在后续工作需要进行分析和探讨。

参 考 文 献

[1] 刘建国, 周涛, 汪秉宏. 个性化推荐系统的研究进展[J]. 自然科学进展, 2009, 19(1):1-15.

[2] Bawden D, Robinson L. The dark side of information: overload, anxiety and other paradoxes and pathologies[J]. Journal of Information Science, 2009, 35(2):180-191.

[3] Goldberg D. Using collaborative filtering to weave an information tapestry[J]. Communications of the Acm, 1992, 35(12):61-70.

[4] Deshpande M, Karypis G. Item-based top- N recommendation algorithms[J]. Acm Transactions on Information Systems, 2004, 22(1):143-177.

[5] Resnick P, Varian H R. Recommender systems[J]. Communications of the Acm, 1997, 40(3):56 - 58.

[6] Hill W, Stead L, Rosenstein M, et al. Recommending and evaluating choices in a virtual community of use[C]// Human Factors in Computing Systems, CHI '95 Conference Proceedings, Denver, Colorado, USA. DBLP, 1995:194-201.

[7] Konstan J A, Miller B N, Maltz D, et al. GroupLens: applying collaborative filtering to Usenet news[J]. Communications of the Acm, 1997, 40(3):77-87.

[8] Lemire D, Maclachlan A. Slope One Predictors for Online Rating-Based Collaborative Filtering[J]. Computer Science, 2007:21-23.

[9] Zhang F, Chang H Y. A collaborative filtering algorithm embedded BP network to ameliorate sparsity issue[C]// International Conference on Machine Learning and Cybernetics, Bonn, Germany. IEEE Xplore, 2005:1839-1844 Vol. 3.

[10] Shardanand U. Social information filtering: algorithms for automating "word of mouth"[C]// Sigchi Conference on Human Factors in Computing Systems, Denver, USA. ACM Press/Addison-Wesley Publishing. 1995:210-217.

[11] Park J, Lee S J, Lee S J, et al. Online Video Recommendation through Tag-Cloud

Aggregation[J]. IEEE Multimedia, 2011, 18(1):78-87.

[12] 徐翔, 王煦法. 协同过滤算法中的相似度优化方法[J]. 计算机工程, 2010, 36(6):52-54.

[13] 尹路通, 于炯, 鲁亮,等. 融合评论分析和隐语义模型的视频推荐算法[J]. 计算机应用, 2015, 35(11):3247-3251.

[14] 杨兴耀, 于炯, 吐尔根&#;依布拉音,等. 考虑项目属性的协同过滤推荐模型[J]. 计算机应用, 2013, 33(11):3062-3066.

[15] Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems : A survey of the state-of-the-art and possible extensions[J]. IEEE Transactions on Knowledge & Data Engineering, 2005, 17(6):734-749.

[16] Pazzani M J, Billsus D. Content-Based Recommendation Systems[C]// Adaptive Web: Methods & Strategies of Web Personalization4321 of Lecture Notes in Computer Science. Springer-Verlag, Braunschweig, German. 2007:325-341.

[17] Pizzato L A, Silvestrini C. Stochastic matching and collaborative filtering to recommend people to people[C]// ACM Conference on Recommender Systems, Recsys 2011, Chicago, USA. 2011:341-344.

[18] Pasquier N, Bastide Y, Taouil R, et al. Discovering Frequent Closed Itemsets for Association Rules[J]. Lecture Notes in Computer Science, 2000, 1540:398--416.

[19] 李瑞敏. 基于社会化网络的个性化音乐推荐算法研究[D].大连：大连理工大学, 2013.

[20] Weng S S, Lin B, Chen W T. Using contextual information and multidimensional approach for recommendation[J]. Expert Systems with Applications An International Journal, 2009, 36(2):1268-1279.

[21] Baeza-Yates R A, Ribeiro-Neto B. Modern Information Retrieval[M]. Beijing, China：China Machine Press, 2011.

[22] Salton G. Automatic text processing[M]. Boston, USA: Addison-Wesley Longman Publishing Co. 1989.

[23] Pazzani M, Billsus D. Learning and Revising User Profiles: The Identification of Interesting Web Sites[J]. Machine Learning, 1997, 27(3):313-331.

[24] Park H S, Yoo J O, Cho S B. A context-aware music recommendation system using fuzzy bayesian networks with utility theory[C]// Fuzzy Systems and Knowledge Discovery, Third International Conference, FSKD 2006, Xi'an, China, Proceedings. 2006:970-979.

[25] Mooney R J, Bennett P N, Roy L. Book Recommending Using Text Categorization with Extracted Information[J]. Recommender Systems Papers from Workshop, 1999:49--54.

[26] Ricci, Francesco. Recommender systems handbook /[M]. USA: Springer, 2011.

[27] Berkovsky S, Eytani Y, Kuflik T, et al. Enhancing privacy and preserving accuracy of a distributed collaborative filtering[C]// Conference on Recommender Systems. New York, USA. 2007:9-16.

[28] Yang X, Guo Y, Liu Y, et al. A survey of collaborative filtering based social recommender systems[J]. Computer Communications, 2014, 41(5):1-10.

[29] Yilmazel B Y, Kaleli C. Robustness analysis of arbitrarily distributed data-based recommendation methods[J]. Expert Systems with Applications An International Journal, 2016, 44(C):217-229.

[30] Wang J, Yin J. Combining user-based and item-based collaborative filtering techniques to improve recommendation diversity[C]// International Conference on Biomedical Engineering and Informatics. Singapore. IEEE, 2013:661-665.

[31] Gavalas D, Konstantopoulos C, Mastakas K, et al. Mobile recommender systems in tourism[J]. Journal of Network & Computer Applications, 2014, 39(1):319 - 333.

[32] 闫祥雨. 基于语义 Web 技术的推荐系统研究[D].太原: 太原理工大学, 2010.

[33] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative filtering recommendation algorithms[C]// International Conference on World Wide Web. Hong Kong. ACM, 2001:285-295.

[34] 邓爱林, 朱扬勇, 施伯乐. 基于项目评分预测的协同过滤推荐算法[J]. 软件学报, 2003, 14(9):1621-1628.

[35] 李雪, 左万利, 赫枫龄,等. 传统 Item-Based 协同过滤推荐算法改进[J]. 计算机研究与发展, 2009(s2).

[36] Chen Y H, George E I. A bayesian model for collaborative filtering[J]. Proceedings of the Seventh International Workshop Artificial Intelligence & Statistics, 1999.

[37] Sarwar B, Karypis G, Konstan J, et al. Application of Dimensionality Reduction in Recommender Systems[J]. In Acm Webkdd Workshop, 2000.

[38] Celeux G, Govaert G. Gaussian parsimonious clustering models[J]. Pattern Recognition, 1995, 28(94):781－793.

[39] Agrawal R, Srikant R. Fast Algorithms for Mining Association Rules in Large Databases[J]. Journal of Computer Science and Technology, 2000, 15(6):619-624.

[40] Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases[C]// ACM SIGMOD International Conference on Management of Data. Philadelphia, USA. ACM, 1999:207-216.

[41] Han J, Pei J, Yin Y. Mining frequent patterns without candidate generation[J]. Acm Sigmod Record, 1970, 29(2):1-12.

[42] Wang J C, Chiu C C. Recommending trusted online auction sellers using social network analysis[J]. Expert Systems with Applications, 2008, 34(3):1666-1679.

[43] Wu X, Zhang Y, Guo J, et al. Web video recommendation and long tail discovering[C]// IEEE International Conference on Multimedia and Expo. Vancover, Canada. IEEE Xplore, 2008:369-372.

[44] Basu C, Hirsh H, Cohen W. Recommendation as classification: using social and content-based information in recommendation[C]// Fifteenth National/tenth Conference on Artificial Intelligence/innovative Applications of Artificial Intelligence. American Association for Artificial Intelligence, Denver, USA 1998:714-720.

[45] Claypool M. Combining Content-Based and Collaborative Filters in an Online Newspaper[C]// Proc. Recommender Systems Workshop at ACM SIGIR. California, USA. 1999.

[46] Pazzani M J. A Framework for Collaborative, Content-Based and Demographic

Filtering[J]. Artificial Intelligence Review, 1999, 13(5):393-408.

[47] Bengio Y. Learning Deep Architectures for AI[J]. Foundations & Trends® in Machine Learning, 2009, 2(1):1-55.

[48] Zhang H P, Yu H K, Xiong D Y, et al. HHMM-based Chinese lexical analyzer ICTCLAS[C]// Sighan Workshop on Chinese Language Processing. Association for Computational Linguistics, Sapporo, Japan. 2003: 758-759.

[49] Durao F, Dolog P. Extending a hybrid tag-based recommender system with personalization[C]// ACM Symposium on Applied Computing. Honolulu, Hawaii. ACM, 2010:1723-1727.

[50] Aldrich J. R. A. Fisher and the Making of Maximum Likelihood 1912-1922[J]. Statistical Science, 1997, 12(3):162-176.

[51] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets.[J]. Neural Computation, 2006, 18(7):1527-1554.

[52] M K. Neural Networks: Tricks of the trade[J]. Canadian Journal of Anaesthesia, 1998, 41(7):658.

[53] Mikolov T, Sutskever I, Chen K, et al. Distributed Representations of Words and Phrases and their Compositionality[J]. Advances in Neural Information Processing Systems, 2013, 26:3111-3119.

[54] Mnih A, Hinton G. A scalable hierarchical distributed language model[C]// Conference on Neural Information Processing Systems, Vancouver, British Columbia, Canada. 2008:1081-1088.

[55] Hinton G E, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing co-adaptation of feature detectors[J]. Computer Science, 2012, 3(4): 212-223.

致　　谢

时间飞逝，光阴如梭，三年的研究生学习生活即将划上句号。回顾在深圳大学求学的这三年时间里，我感触颇深。我遇到了良师，汲取了知识，历练了自己，收获了友情，我对这段美好的时光充满感激！

首先，要特别感谢的是我的导师李霞教授。衷心感谢李老师对我这三年多时间的指导与栽培。学业上，李老师为我指明方向，使我的学术能力得到一定的提升，让我感受到沉浸在学术世界的快乐。同时，李老师严谨的学术态度、务实的学术精神以及独特的学术研究视角深深地感染了我，这将对我今后的学习、生活和工作带来深远的影响。此外，本论文从选题到成稿的整个过程都离不开李老师的悉心指导，在此向李老师致以学生最诚挚的谢意！

其次，我要感谢我的副导师陈亮老师。陈老师从我论文的选题到最后论文的写作给了我许多的帮助和建议。学业上，在我困惑时陈老师给我答疑解惑，使我在科研的道路上坚定前行。生活上，陈老师对我悉心照顾，为我排忧解难。谨在此表示我衷心的感谢。

然后，我要感谢我的同学同学，你们陪伴着我一路成长，让我的研究生生活充满了友爱与欢乐。感谢各位学弟学妹们在学习与工作中对我的支持与协助，感谢我的室友在这三年中给予我的支持与鼓励，感谢你们让我获得了珍贵美好的友情。

再者，我要深深的感谢我的家人。是你们的支持，让我开始这段学术生涯；是你们的鼓励，让我顺利走完这段学术生涯；是你们的爱，让我坚持到了最后。谢谢你们给予我的支持和鼓励，谢谢你们给予我的爱和期望！祝我的父亲母亲身体健康！幸福快乐！

最后，特别感谢各位论文评审专家教授，感谢你们在百忙之中抽出时间审阅我的论文。由于本人水平有限，文中难免有疏漏和错误之处，请各位专家教授批评指正！

攻读硕士学位期间的研究成果

[1] 陈亮、高睿、王娜、李霞；跌倒检测方法、系统及基于该系统的跌倒自动报警器；中国；申请号或者专利号：201510597866.X 已受理