

清 华 大 学

# 综 合 论 文 训 练

题目：基于跨空间解耦表征的社会化  
推荐研究

系 别：计算机科学与技术系

专 业：计算机科学与技术

姓 名：何昕雨

指导教师：王 鑫 助理研究员

2022 年 6 月 13 日

## 关于学位论文使用授权的说明

本人完全了解清华大学有关保留、使用学位论文的规定，即：学校有权保留学位论文的复印件，允许该论文被查阅和借阅；学校可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存该论文。

(涉密的学位论文在解密后应遵守此规定)

签 名： 何昕雨 导师签名： 王鑫 日 期： 2022/6/13

## 中文摘要

这些年来，互联网上的用户数飞速增长，由此催生出了诸如电商、短视频平台、线上电影、电子图书馆等应用。与此同时，如何设计推荐系统来利用用户反馈的偏好信息为用户推荐他们可能感兴趣的内容、优化用户体验也成为了一个热点问题。在互联网兴起的同时，社交网络的飞速发展传统推荐任务带来了新的思路。在对人们行为的观察中我们可以发现，用户的选择倾向会受到身边好友的影响，比如我们被朋友“种草”某样东西就是其中一个例子。所以，通过对社交网络的分析，我们可以更好地理解人与人之间的行为、关系，进而帮助我们为个人推荐更合适的物品。

然而，目前社会化推荐系统学习所得到的物品表征和用户表征的可解释性都比较差，学习得到的表征中不同影响因子对应的特征都完全融合在一起。为了解决这一问题，本研究首先通过解耦表征学习来找到物品或用户特征中独立的因子，将原本耦合在一起的特征解耦，让最终得到的解耦表征可以被分离为若干个影响因子的对应特征，提高模型的可解释性以及可操作性；同时从注意力机制中获得灵感，进一步利用解耦的图神经网络对社交网络进行解耦分析，以提高推荐的准确性。

最终本文实现了一套解耦的社会化推荐系统，实验表明该系统在数据集上比没有利用社交信息进行优化的解耦推荐系统的召回率 Recall 和归一化折损累计值 NDCG 有明显提升。

**关键词：**社会化推荐；解耦表征；图神经网络；推荐系统

## ABSTRACT

In recent years, the volume of Internet has increased rapidly, which boosted the appearance of various online platforms such as e-commerce, short video platform, online film, e-library and so on. In the meantime, how to design recommender systems to exploit users' feedback to recommend the content they may be interested in and optimize users' experience has also become a hot topic. The rapid development of social networks has brought new methods to the traditional recommendation task. In the observation of people's behavior, we can find that users' tendency are normally being affected by their friends. For instance, we have tends to buy things that our friends recommend to us serves as an example. Therefore, by analyzing social networks' features, we can better understand the behavior and relationship between people, and so far helps us to recommend more suitable items for individuals.

However, the interpretability of item representations and user representations studied by current social recommendation systems is relatively poor. The representations corresponding to different influencing factors are fully entangled together in the trained item representations and user representations. In order to solve this problem, our work first disentangles the independent factors in the recommendations of items and users through disentangled representation learning to improve the interpretability of the model; At the same time, inspired by the residual network, disentangled graph neural networks are used to exploit the social networks, improving the efficiency of recommendation results.

All in all, this paper implements a disentangled social recommendation system. The experiments show that the Recall rate and Normalized Discounted cumulative gain (NDCG) of the system outperforms those disentangled recommendation systems without social information optimization.

**Keywords:** social recommendation; disentangled representations; Graph Neural Networks; recommender systems

# 目 录

第 1 章 引言 .....	1
1.1 研究背景 .....	1
1.2 当前方法的问题 .....	2
1.3 解耦的社会化推荐系统 .....	3
1.4 论文结构 .....	3
第 2 章 相关工作 .....	5
2.1 推荐系统 .....	5
2.1.1 基于神经网络的推荐系统 .....	6
2.1.2 社会化推荐系统 .....	7
2.1.3 解耦的推荐系统 .....	8
2.2 解耦的图神经网络 .....	9
2.3 注意力机制 .....	10
第 3 章 解耦的社会化推荐系统 .....	11
3.1 问题定义 .....	12
3.2 初步模型结构 .....	12
3.2.1 聚类模块 .....	12
3.2.2 编码模块 .....	12
3.2.3 解码模块 .....	14
3.3 目标函数 .....	14
3.4 注意力机制优化 .....	14
3.4.1 社交信息解耦 .....	15
3.4.2 注意力机制 .....	16
第 4 章 实验结果 .....	17
4.1 数据集 .....	17
4.2 实现细节 .....	17
4.3 实验结果 .....	18

第 5 章 结论 .....	20
插图索引 .....	21
表格索引 .....	22
参考文献 .....	23
致 谢 .....	25
声 明 .....	26
附录 A 外文资料的书面翻译 .....	27
在学期间参加课题的研究成果 .....	58

## 主要符号表

GNN	图神经网络 (Graph Neural Network)
GCN	图卷积网络 (Graph Convolutional Network)
dropout	随机失活
Tri-training	三体训练法
VAE	变分自编码器 (Variational Autoencoder)
Softmax	柔性最大传递函数
Gumbel-Softmax	坎贝尔柔性最大传递函数
Linear	神经网络线性层
diag	对角矩阵
$\mathcal{N}$	正态分布
$D_{KL}$	Kullback-Leibler 差异, 也称作相对熵 (Kullback-Leibler Divergence)
sigmoid	S 型生长曲线
Recall@K	仅考虑前 K 项的召回率 (Recall at top K)
NDCG@K	仅考虑前 K 项的归一化折损累计增益 (Normalize Discounted Cumulative Gain at top K)

# 第 1 章 引言

## 1.1 研究背景

21 世纪以来互联网飞速发展，各类电子应用也在网络上逐渐发展、成熟，这成为了推荐技术诞生、发展的源动力。如今的一个网站或应用，如淘宝、豆瓣、各大视频网站等等，都可以从用户行为和用户反馈中很轻松地获得大量地用户信息。以一个购物平台为例，它可以从用户是否点击浏览了某个商品、是否选择购买了某个商品、是否选择评价了该商品以及显式地在评价时给予该商品的分值等多个维度来获得该用户对于该商品以至于同类相似商品的购买倾向。有此丰富的数据来源为基础，平台就可以利用这些信息来分析提取出用户的需求和偏好，在实现首页推荐、搜索前排推荐等网页/应用常见的功能时，就可以借助于这些分析提取出的高阶特征来做个性化的推荐。这不仅能优化用户体验，同时也能吸引用户继续使用该平台，提高用户粘性，因此现在我们知道的很多包括淘宝、豆瓣、抖音等应用背后都有很多推荐算法的支持。

为了进一步提升推荐效果，研究人员又直观地观察到了这样一个现象：社交关系也会影响人们对于物品的偏好和选择。例如，微博抖音上的网红博主会进行产品推荐，一部分博主粉丝就会被“种草”；我看到朋友买的衣服很好看，我想买衣服的时候可能会买同款或类似的款式。总结而言，基于“如果我的朋友喜欢/买了某样产品，那么我也可能会喜欢/购买某样产品”这样一个假设，研究人员提出了**社会化推荐系统**这一新的研究方向。这里的“朋友”不仅仅局限于社交关系上的朋友，而是类似于“信任度”，比如我关注了某个博主就是这样一种“信任度”的体现。

解耦表征学习则是近些年来新兴的一个机器学习领域。可解释性(interpretability)是人工智能领域中一个非常重要的概念，一般来讲可以被分为数据的可解释性以及模型的可解释性。前者指让模型中的数据可以更易于被人理解，后者是指让模型的运作机制和有效性可以被解释。它一方面是对人工智能安全的保障，另一方面也使得开发者对模型黑盒的运行有更大的掌控力。**解耦表征学习就是为了提高以学习表征为目标的模型的数据可解释性而出现的**。假设现实数据是由一系列可解释的因子  $k_1, k_2, \dots, k_K$  所决定的，即现实数据  $y$  可以表示为一个未知的复杂函数  $f(k_1, k_2, \dots, k_K)$ 。传统的表征学习目标是得到一个表征  $z$ ，它能在抽象地表示



我们所需要的现实数据  $y$  中所包含的各种信息。那么解耦表征学习的目标就是使得最终学习得到的表征  $z$  可以被分离为  $z_1, z_2, \dots, z_K$ ，每个子表征对应于一个影响因子，使得模型最终得到的表征可以更利于人们的理解。仍以购物平台为例，在以用户购买过的商品为数据的推荐任务中，我们的现实数据就是用户购买过的商品列表，影响决定用户购买这些物品的因子可能包括一些高层次的因子，如用户需要的物品类别；也可以被进一步分解为一些更底层的影响因子，比如用户喜欢的颜色、用户喜欢的风格等等。在这样一个推荐系统中，解耦表征学习最终得到一个可以被分离为  $z_1, z_2, \dots, z_K$  的表征  $z$ ， $z_1$  中解耦分离得到了对应于决定用户是否喜欢购买游戏装备的特征， $z_2$  中解耦分离得到了对应于决定用户是否喜欢网购裙子的特征，...，等等。在完全解耦的条件下，每个子特征  $z_k$  中的每个维度应当都可以被分别解耦表示对应于某一个底层影响因子，比如  $z_2$  中的某个维度可能代表了用户喜欢购买的衣服颜色。

本文的工作在社会化推荐系统和解耦表征学习的交叉领域中展开，希望用解耦表征学习来提升社会化推荐系统的可解释性。

## 1.2 当前方法的问题

社会化推荐系统的实现目前大概可以分为以下几类：

- 将社交信息作为一种背景信息视为用户的一种特殊特征，利用情境感知推荐系统（Context-aware Recommender Systems）的方法来做推荐。
- 以社交网络为中心来做推荐。这种情况需要已知一个社交网络，再基于前文中提到的假设来推断某个用户的偏好。
- 通过用户与社交网络交互过程中反馈的标签或者评论来协助进行分类/分析。

以上这几类方法分别针对不同场景、不同数据采用了不同方法处理推荐任务。而在本文工作中，我们处理的主要是第二种场景，即在已知社交网络的情况下进行社会化推荐。

目前以社交网络为中心的社会化推荐算法中，主流方法还是通过学习物品表征和用户表征来实现的。然而这些模型的可解释性却比较差，最终得到的表征中对应于各个因子的子表征完全耦合在一起，因此如何将社会化推荐系统解耦，提高社会化推荐系统的可解释性，也成为了一个重要问题。图神经网络是常用的对社交网络进行训练分析的方法之一，因此本文希望可以借助图神经网络来学习社交信息实现社会化推荐系统并解耦。事实上，目前已经有一些前人对解耦推荐

系统和解耦图神经网络进行了一些探索和研究，如 Ma et al.<sup>[1]</sup>和 DisenGCN<sup>[2]</sup>。因此，我们主要的研究难点就在于如何设计一个可以被解耦的社会化推荐系统，在发挥社交信息对推荐的辅助作用的同时，让模型具有一定的可解释性。

### 1.3 解耦的社会化推荐系统

考虑到在不同应用场景下社交信息对于用户选择的影响程度不同，例如朋友之间互相推荐的影响对于用户对电影的选择的影响程度可能大于用户对于日用品选择的影响程度。因此在注意力机制<sup>[3]</sup>的启发下，我们增加了参数  $\gamma$ ，利用解耦的图神经网络分析社交信息，与此同时将先前根据用户历史购买信息提取出的特征信息和得到的社会化用户特征按照参数进行比例结合，以在获得带有社交信息的用户特征的同时，保留部分从原始数据中提取出的用户特征。同时，如果仅仅采用社会化推荐后的结果，可能导致朋友之间的特征趋同性过高，导致推荐中个性化的部分大大减少。这并不是我们所想要的，因此用注意力机制的方式对社交信息对最终结果的影响比例进行一定的控制。

本文以 Ma et al.<sup>[1]</sup>中对于推荐系统解耦的分析为基础提出了一种解耦的社会化推荐系统。Ma et al.<sup>[1]</sup>在将推荐系统解耦上的成就让本文可以遵循该工作的大致思路来实现社会化推荐系统的解耦，而 DisenGCN<sup>[2]</sup>在图神经网络解耦上的成功则为本文解耦社会信息提供了基础。在设计社会化推荐系统的过程中，本文也进行了不同的尝试，最终提出了以解耦的特征做初始化的解耦图神经网络块，采用类似于注意力机制的方式让最终得到的解耦特征能更有效地反映用户偏好。

### 1.4 论文结构

本文后续内容的结构组织如下：

第二章中，我们介绍了与本文提出的模型相关的一些工作，包括基于神经网络的推荐系统、社会化推荐系统、解耦的推荐系统、解耦图神经网络、注意力机制。

第三章中，我们介绍了不使用注意力机制优化的解耦的社会化推荐系统的网络架构和训练过程，以及参考注意力机制进行优化后的社会化推荐系统的网络架构和训练过程。

第四章中，我们介绍了实验细节并给出了一些实验结果，做了简要分析。

第五章中，给出了对本文的结论与总结，并提出了一些仍然存在的不足之处

以及未来可能的发展方向。

## 第 2 章 相关工作

在本章中首先介绍了推荐系统领域的发展状况，并对该领域中与本职工作相关的两个子领域——社会化推荐系统和解耦的推荐系统进行进一步展开探讨，以期让读者本文所探讨研究的问题背景有一个大致了解。

接着，本章中也简要介绍了本职工作所涉及的图神经网络和注意力机制，并对尤其是解耦的图神经网络相关工作进行了更详尽的说明。

### 2.1 推荐系统

如前文所述，推荐系统诞生于飞速发展的互联网及互联网应用中，其目标可以简要概括为利用收集到的各类信息来推断用户的对某些物品的偏好。互联网上可以被收集到的用户信息可以被分为两种，一种是显式的反馈，比如说 Netflix、豆瓣要求用户给电影打分，或者选择是否对其感兴趣；另一种是比较隐蔽但是更便于收集的数据，例如是否点击浏览/购买了某个物品。通过分析用户和物品之间历史中的交互关系，就可以对用户的未来偏好和选择进行推断。当然这不是在所有条件下都成立，对于一个靠用户即时指定要求的系统来说，这种依靠历史观察来推断未来可能的系统就不适用了。但是在现实中大部分应用场景下都有一定时间上的一致性，可以采用从过去中观察规律并推至未来的方法，因此推荐系统在现实中有大面积的应用。

推荐系统的实现方法可以被分为大致以下几种思路：

- **基于相似邻居做推荐的方法**：该类方法可以从用户相似度或者物品相似度两个方向分别展开，基本思想就是利用一定的评价指标来从历史数据中衡量两个用户或者两个物品之间的相似性，找到和该用户/物品相似的“邻居”，然后将相似用户感兴趣的商品推荐给该用户，或者将相似的物品推荐给购买过该商品的人。衡量相似性的指标就要视不同方法而定了，但一般历史浏览/购买/打分记录越相似的用户相似性的指标值就越高，历史中被浏览/购买/打分的记录越相似的物品相似性的指标值就越高。
- **基于模型做推荐的方法**：该类方法用处理预测任务的模型来从数据中学习参数，预测未来用户的偏好。其中一种比较常见的方式是学习用户和物品的特征向量，并用向量之间的一个**打分函数**来估计用户选择该物品的倾向。这一

类方法即使是面对比较稀疏的用户和物品之间的历史关系数据也能做到比较好的处理。

- 基于情境做推荐的方法：该类方法适用于物品本身带有一些标签的情况，将用户分开独立考虑，用该用户对其他物品的偏好以及这些物品之间特征上的相似性来判断用户对该物品的偏好。

本文的工作是以基于模型做推荐的方法为背景展开的，该类方法的优势在于训练过程是可以提前完成的，后续进行预测时只需要使用预训练好的模型即可。并且和其他针对于某个用户-物品对进行处理的方法不同，该方法在训练过程中利用了整个数据集，可以挖掘出数据整体背后的一些模式特征，因此使用也很广泛。具体使用到的模型可能包括决策树、贝叶斯分类器、回归模型、支持向量机、神经网络等等。

### 2.1.1 基于神经网络的推荐系统

近些年来由于神经网络的迅猛发展，涌现出了很多利用神经网络模型来实现推荐系统的研究。

神经网络本身就像一个黑盒子，我们不需要根据数据本身的特点和结构来精心设计相应的算法，只要数据能够符合黑盒子的输入输出要求就可以用神经网络来训练，根据具体目标设计损失函数和网络结构即可。一种常用的数据处理方法是将信息转变为“01”单值，丢失或不知道的信息位置被置为0，有确定信息的位置被置为1。这种单值的数据格式适用于很多现实中的实验数据，以购物平台的用户购买历史数据为例：假设有  $N$  位用户， $M$  件商品，则输入数据可以构造为一个  $N * M$  的稀疏矩阵  $X$ ，其中  $x(u, i) = 1$  表示用户  $u$  是否购买过物品  $i$ ， $x(u, i) = 0$  则表示在当前的记录中用户  $u$  没有购买过商品  $i$ 。事实上，用户、物品之间的这种关系可以被看作是一个二分图（bipartite graph），这样的图中的节点可以被分为两个独立的集合，集合内部节点之间没有边相连，属于不同集合的节点之间可以有边相连。而上文提到的这种数据处理方法实际上就是无向二分图的邻接矩阵。

最新的一些基于神经网络的推荐系统的研究包括：

NGCF<sup>[4]</sup> 直接在用户-物品二分图上采用图神经网络（Graph Neural Networks, GNN）学习用户与物品表征，以获取用户-物品关系中的信息，而不是仅仅从用户、物品的标签或特征来学习用户、物品表征。为此，该工作提出了 NGCF 框架，为其设计了处理用户-物品二分图的嵌入传播层，实现了有效的在用户-物品二分图上的消息传递机制。

LightGCN<sup>[5]</sup>指出目前从 GCN (Graph Convolutional Networks, 图卷积网络) 中获得灵感的推荐系统工作 (比如 NGCF<sup>[4]</sup>) 中的特征转换和非线性激活部分对于改善推荐效果没有任何帮助, 并且会导致训练的难度增加。因此, LightGCN 只保留了在用户-物品二分图上进行邻居聚合的部分, 并得到了比 NGCF 高出 16% 左右的推荐效果。

SimGCL<sup>[6]</sup>的工作则是基于对比学习的, 它通过实验指出以往基于对比学习的推荐系统使用的图增强的思想其实并不重要, 对比学习实际上是通过学习更均匀分布的用户/物品表征来隐式消除偏差的。因此该模型撇弃了图增强的方法, 转而改为在嵌入空间中加入均匀噪声来构造对比。

SGL<sup>[7]</sup>提出了基于 GCN 方法的两个问题: 度数高的节点对于表征有更大的影响, 导致度数小的节点 (在图中占比更大) 的表征被污染; 由于邻居聚合导致边的影响增大, 所以也导致这些边中可能出现的噪声对表征的影响增大, 让模型对噪声变得更加敏感。为了提升模型的准确度和鲁棒性, SGL 以 LightGCN 为基础, 采用随机丢弃边 (edge dropout)、随即丢弃节点 (node dropout) 和随机游走 (random walk) 三种方法为某一个节点创建不同场景, 让该节点在不同场景下能一致地与其他节点区分开。该工作也从理论上证明了这种方法可以自动的寻找到强负例。

### 2.1.2 社会化推荐系统

在引言中已经对社会化推荐系统的几类实现方法进行了大致的介绍, 本工作是以以社交网络为中心的社会化推荐系统为背景展开的。社会化推荐系统在某些应用场景下之所以相较于一般推荐系统能有准确度上的提升, 是基于用户和用户在社交网络的朋友的偏好会比较相似这一假设。举个简单的例子, 用户在首次购买某类商品时如果自己选择下单都会有试错成本, 但是如果用户看到朋友用的一款效果很好或者很推荐, 那么他就很有可能购买同款以避免试错。

以社交网络为中心的社会化推荐系统的具体设计思路又有很多, 以下仅列出几篇最新的研究以供读者理解该方向的最新进展:

SEPT<sup>[8]</sup>是一篇基于自监督学习的工作, 它认为当前基于自监督学习的方法只在两个不同场景的节点之间建立了双向映射, 忽略了来自其他节点的自监督信息。因此它采取了 Tri-training 的训练思路, 在多场景编码的背景下, 在增强场景中设计了三个图编码器, 其中一个用于做推荐, 另外两个用于获取来自其他用户的自监督信息, 并用这些自监督信息来更新这三个编码器。

MHCN<sup>[9]</sup>认为用户之间的社交关系是非常复杂的, 用户之间的社交联系可以

是高维的。超图可以被用于很好地建模用户之间这种复杂的高维关系，但是目前的社会化推荐系统还没有可以处理超图的工作，因此该论文提出了一种用超图来改进社会化推荐效果的模型。该模型是一个多通道超图卷积网络，每个通道通过超图卷积来编码一个表示了某种用户之间高维关系的超图，并最后将各个通道聚合在一起。同时，该模型还引入了超图中的自监督方法，来保留不同类型的用户之间高维关系的固有特征。

DiffNet<sup>[10]</sup>指出目前的社会化推荐系统都只是利用本地邻居来进行更新的静态模型，没有模拟全局社交网络中的递归扩散过程，但事实上用户的偏好是会随着社会作用在社交网络上传播和扩散的递归过程发生变化的。为此 DiffNet 提出了一个以逐层作用传播结构为基础的深度作用传播网络来模拟用户在递归社交扩散过程中是如何被影响的。

### 2.1.3 解耦的推荐系统

目前在解耦推荐系统领域的研究还比较少，有待后人探索。Ma et al.<sup>[1]</sup>是近几年一篇比较经典的解耦推荐系统的论文，本工作也以其为背景开展。

Ma et al.<sup>[1]</sup>认为用户的行为是由用户做决定过程中的很多因素进行复杂交互所决定的。它提出了宏观解耦和微观解耦两个相对应的概念，宏观解耦将与用户兴趣和意图相关联的高层次概念（如物品类别）进行解耦，微观解耦则将用户对于某个物品更细微的方面（如用户喜欢的颜色、需要的大小）的偏好进行解耦。该方法首先推断与用户意图相关的高层次概念，然后获取用户对于这些高层次概念的偏好。接着用一个源于信息论解释的 VAE 的微观解耦正则器来让表征的每一个维度都能独立地反应一个单独的低层次概念。为了实现微观解耦，该模型根据 VAE 进行改进设计了如下的训练目标：

$$E_{p_{\theta}(C)}[E_{q_{\theta}(z_u|x_u, C)}[\ln(p_{\theta}(x_u|z_u, C))]] - \beta \cdot D_{KL}(q_{\theta}(z_u|x_u, C)||p_{\theta}(z_u))$$

并理论证明了该训练目标可以实现微观解耦。该工作在实验上超过了以往的解耦推荐系统的推荐准确度，并且在实验上也确实达到了一定的解耦效果。

除了这一篇以外，还有以下几篇工作也是与解耦推荐系统相关的：

DESTINE<sup>[11]</sup>的工作是用于解决广告中的点击通过率预测（Click-Through Rate Prediction, CTR）问题的，但是也可以应用于很多推荐问题中，即预测某物品有多大概率被用户点击。该工作通过解耦的自注意框架将一元特征重要性与二元特征交互解耦，一元学习的部分建模某个特征相对于其他特征的重要性，二元交互学习的部分学习某个特征对的影响。这篇论文主要是实现了一阶项和二阶项的解耦，



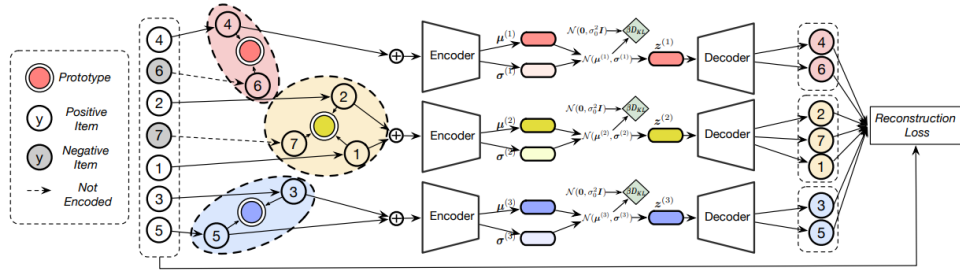


图 2.1 Ma et al.<sup>[1]</sup> 的模型结构。原图来自论文<sup>[1]</sup>。

和本工作中所要实现的解耦还不太一样。

DGCF<sup>[12]</sup>是基于用户-物品二分图和解耦图神经网络所实现的推荐系统。它将不同的用户意图分别在胶囊网络（capsule networks）的不同胶囊中用动态路由和图神经网络（GNN）中的嵌入传播机制学习用户-物品连接和高阶连接中的信息，并用独立建模的距离来保证不同用户意图之间的独立性。和 Ma et al.<sup>[1]</sup>类似地，该方法也是从用户角度实现的解耦。

DIR<sup>[13]</sup>则是实现了物品表征的解耦，将每个物品表征用几个对应于不同特征的特征向量组合表示。其中特征又包含显式特征和隐式特征，显式特征指数据集里明确标记了的物品特征，隐式特征则是待学习的物品特征。该工作提出了 LearnDIR 算法，每轮将隐式特征分配给某个物品，然后根据使用的具体方法进行学习并优化向量。它从实验上证明了 LearnDIR 算法在矩阵分解（Matrix Factorization）和递归神经网络（Recurrent Neural Networks）两类方法上是有效的。

## 2.2 解耦的图神经网络

图这种数据结构由于其相对于一般欧式空间向量的强大表示能力而一直受到科研人员的广泛关注。因此，在神经网络崛起并在许多线性向量数据得到成功应用后，人们提出了处理图数据的神经网络，即图神经网络 GNN。自然而然地，在研究解耦机器学习的过程中，如何解耦图神经网络也成为研究人员关心的问题之一。

该方向比较经典的一篇论文是 DisenGCN<sup>[2]</sup>。该工作创新地提出了一种邻居路由的方法，该方法可以动态地识别导致某节点与其邻居间有边的潜在因素，然后将对应于不同因素的邻居分配到该因素的管道中进行分开处理，来提取和卷积对应于某一个因素的特征信息。



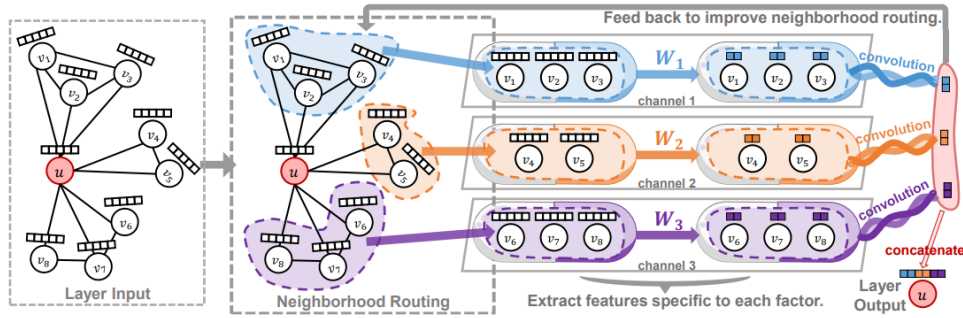


图 2.2 DisenGCN 的模型结构。原图来自论文[2]。

## 2.3 注意力机制

注意力机制最初是为了解决在自然语言处理中语句中靠前的词的信息会被丢失的问题，而引入注意力机制来将“注意力”放在当前预测词的上下文中有用的词上。在注意力机制中，一个上下文向量的计算方法为：

$$c_i = \sum_{j=1}^T \alpha_{i,j} h_j$$

其中  $\alpha_{i,j}$  是权重，需要由神经网络计算得到。但是由于本工作只是从注意力机制这种加权的方法中获得了一些改进思路，所以就不对权重的具体训练方法进行展开了。

### 第 3 章 解耦的社会化推荐系统

本工作首先初步实现了一个解耦的社会化推荐系统，后续用注意力机制做了进一步的优化，初步模型的大致结构如图 3.1。

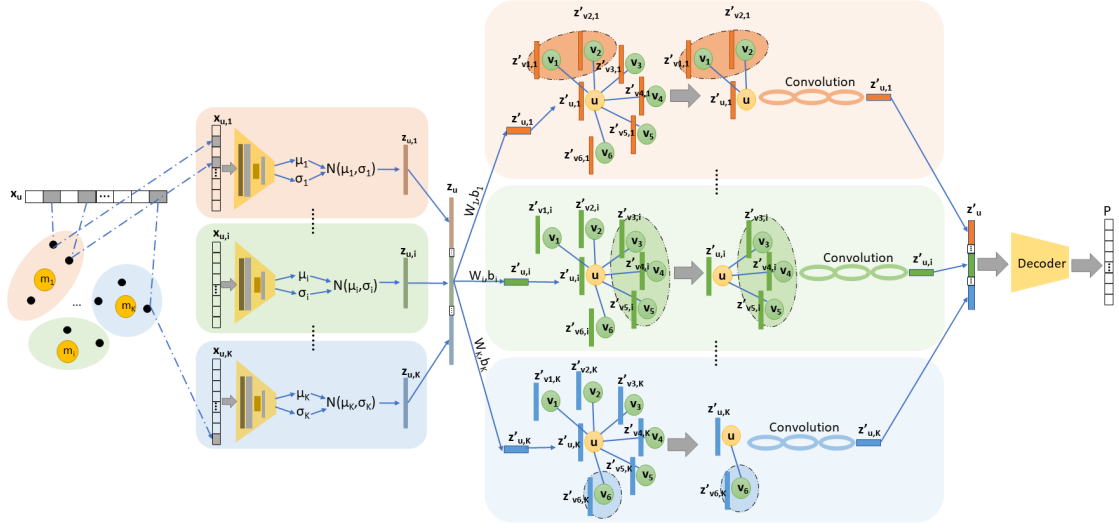


图 3.1 初步模型的整体结构。

可以看到整体模型和 Ma et al.<sup>[1]</sup>保持了较高度的一致，总体可以分为聚类、编码、解码三个模块，其中编码模块又可以被分为压缩编码和路由传播两个部分。模型的整体思路是首先在聚类模块中将物品表征以不同因子的表征为中心进行聚类，以估计某个物品是否从属于某个因子。接着在编码模块中分为  $K$  个通道，每个通道用于训练提取对应于某一个高阶因子的表征。在每个通道中先将用户购买的那些物品中与该因子相关的部分提取出来，然后经过编码压缩为指定长度的隐层表征，该表征通过 DisenGCN<sup>[2]</sup>的路由机制来学习社交网络中朋友的信息。最后将各个通道学习到的表征拼接在一起得到最终的用户表征，并经过解码器给出一个预测用户是否会购买/点击某物品的概率表，并可以进一步处理结果给出一个推荐排序。

本章将从问题定义、初步模型结构、目标函数、注意力机制优化几个部分进行展开。

### 3.1 问题定义

本工作所设计的社会化推荐系统的输入数据需要用户-物品交互历史和用户社交网络两个部分。对于包含  $N$  个用户和  $M$  个物品的推荐系统来说，用户-物品交互历史以一个  $N * M$  的单值矩阵  $X$  表示，其中  $x_{u,i} = 1$  表示第  $u$  个用户购买过物品  $i$ ，否则  $x_{u,i} = 0$ ；用户社交网络则以无权图的形式呈现，用户  $u$  和用户  $v$  之间有边则代表这两个用户之间存在社交关系。

给定输入数据，本工作的目标是学习得到解耦的用户表征  $\{z_u\}_{u=1}^N$ 。在给定  $K$  的情况下，可以将用户表征解耦成对应于  $K$  个高阶因子的子表征  $z_u = [z_{u,1}, z_{u,2}, \dots, z_{u,K}]$ ，其中用户表征为  $d'$  维，每个子表征为  $d$  维， $d' = Kd$ 。同时希望子表征尽量实现微观解耦，并且可以反应一定社交关系对用户偏好的影响。

最终，通过在解码器中计算用户表征和物品表征的相似度，我们可以给出一个推荐分值或推荐列表。

### 3.2 初步模型结构

在这一节中将对初步社会化推荐系统模型中的各模块进行分开介绍。

#### 3.2.1 聚类模块

该模块与 Ma et al.<sup>[1]</sup>中的聚类方法保持一致，将待学习的物品表征  $\{h_i\}_{i=1}^M$  围绕待学习的高阶因子表征  $\{m_k\}_{k=1}^K$  进行聚类。在本模型中使用余弦相似度  $s_{i,k}$  来估计某物品  $i$  属于某高阶因子所代表类别  $k$  的概率，用一个超参数  $\tau$  来控制相似度的取值范围，然后用耿贝尔柔性最大传递函数（GUMBEL-SOFTMAX）来使得最终得到的分类向量逼近于一个独热向量（one-hot 向量）。具体公式如下：

$$s_{i,k} \leftarrow h_i^T m_k / (\tau \cdot \|h_i\|_2 \cdot \|m_k\|_2), k = 1, 2, \dots, K.$$

$$c_i \sim \text{GUMBEL-SOFTMAX}([s_{i,1}; s_{i,2}; \dots; s_{i,K}])$$

$c_i$  即为得到的代表物品分类的近似 one-hot 向量，其中  $c_{i,k}$  代表物品  $i$  是否属于类别  $k$ 。

#### 3.2.2 编码模块

该模块分为  $K$  的通道，每个通道学习对应于某一个高阶因子的子表征。在第  $k$  个通道从某个用户的购买历史信息  $x_u$  中提取表征之前，需要将其用户购买信息中对应于不同高阶因子的特征解耦出来，也即筛选出用户购买过的商品中属于类

别  $k$  的商品, 得到只包含与第  $k$  个因子相关的用户购买信息  $x_{u,k}$ 。这就需要用到在聚类模块学习得到的物品类别 one-hot 向量:

$$x_{u,k,i} = x_{u,i} * c_{i,k}$$

在得到  $x_{u,k}$  之后就可以将该特征向量压缩并从中提取隐层的子表征, 即编码模块中的压缩编码部分。我们假设对应于第  $k$  个高阶因子的子表征是从正态分布  $\mathcal{N}(\mu_{u,k}, \text{diag}(\sigma_{u,k}^2))$  中采样得到的, 因此我们在编码部分的目标就是学习得到  $\mu_{u,k}$  和  $\sigma_{u,k}$ 。在本工作的模型实现中我们使用了两个 Linear 线性层来实现初步的特征压缩和表征提取, 第一个 Linear 层将高维的用户特征压缩为指定  $d$  维的隐层表征, 第二个 Linear 层将  $d$  维的表征编码为  $2d$  维的表征, 前  $d$  维  $a_{u,k}$  含有  $\mu_{u,k}$  的信息, 后  $d$  维  $b_{u,k}$  含有  $\sigma_{u,k}$  的信息。然后通过从学习得到的正态分布中采样得到初步的子表征:

$$\mu_{u,k} \leftarrow a_{u,k} / \|a_{u,k}\|_2, \sigma_{u,k} \leftarrow \sigma_0 \cdot \exp(-\frac{1}{2} b_{u,k})$$

$$z_{u,k} = \mu_{u,k} + \epsilon \cdot \sigma_{u,k}, \epsilon \sim \mathcal{N}(0, 1)$$

其中  $\sigma_0$  是超参数, 我们认为整体表征空间是从  $\mathcal{N}(0, \text{diag}(\sigma_0^2))$  采样得到的。 $z_{u,k}$  就是压缩编码部分最终得到的用户  $u$  对应于第  $k$  个高阶因子的子表征。

在路由传播部分, 我们遵循了 DisenGCN<sup>[2]</sup> 中的邻居路由机制。每次训练在社交网络无权图上进行  $T$  轮路由传播, 每轮根据一些原则用用户邻居的子表征更新用户的子表征。为了将用户表征解耦, 我们在图上进行路由时也需要考虑到有哪些邻居和用户之间更可能在高阶因子  $k$  上互相影响。一个很简洁的解决方法是用用户和邻居在第  $k$  个因子上的子表征之间的相似度来衡量该邻居有多大可能在因子  $k$  上对用户产生影响, 对于  $u$  的每一个邻居  $v$ :

$$p_{v,k} \leftarrow z_{v,k}^T y_k / \tau$$

$$p_{v,k} \leftarrow \text{softmax}(p_{v,k})$$

其中  $y_k = z_{u,k}$ ,  $p_{v,k}$  即为我们认为用户  $v$  有多大可能在因子  $k$  上对用户  $u$  产生影响。然后根据该概率来用邻居的子表征更新用户子表征, 由于在这个过程中概率会随着用户表征的变化发生变化, 所以需要进行若干  $T$  轮路由更新, 即上面计算概率和下面路由更新的步骤需要重复进行  $T$  次:

$$y_k \leftarrow z_{u,k} + \sum_{v:(u,v) \in G} p_{v,k} z_{v,k}$$

$$y_k \leftarrow y_k / \|y_k\|_2$$

在进行  $T$  轮路由传播后, 得到的  $y_k$  即为该用户在第  $k$  个因子上学习到了一定社交信息的子表征, 将其重新赋回给  $z_{u,k}$ :  $z_{u,k} = y_k$ 。

### 3.2.3 解码模块

解码器比较简单，根据物品的类别  $k$ ，判断用户在第  $k$  个因子上的子表征与物品表征的相似度来衡量用户未来购买/点击该商品的概率。

$$p_{u,i} = \sum_{k=1}^K c_{i,k} \cdot \exp(z_{u,k}^T h_i / (\tau \cdot \|z_{u,k}\|_2 \cdot \|h_i\|_2))$$

$p_{u,i}$  越大则越应该将物品  $i$  推荐给用户  $u$ ，所以将  $\{p_{u,i}\}_{i=1}^M$  从大到小排序就可以得到一个推荐列表，用于实际用户推荐。

### 3.3 目标函数

目标函数沿用了 Ma et al.<sup>[1]</sup> 中的目标函数，以实现宏观解耦和微观解耦，具体目标函数形式如下：

$$E_{p_\theta(C)}[E_{q_\theta(z_u|x_u, C)}[\ln(p_\theta(x_u|z_u, C))]] - \beta \cdot D_{KL}(q_\theta(z_u|x_u, C) || p_\theta(z_u))$$

可以看到该目标函数分为两个部分，左半部分近似于：

$$\sum_{i: x_{u,i}=+1} \ln p_{u,i}$$

右半部分近似于：

$$D_{KL}(\mathcal{N}(\mu_u, \text{diag}(\sigma_u)) || \mathcal{N}(0, \sigma_0 \cdot I))$$

$\beta$  则是一个超参数，用于在两部分损失函数之间根据具体需求进行协调平衡。左半部分保证学习得到的表征重构得到的推荐概率接近于真实数据，右半部分则保证各因子的子表征尽量解耦，利用 KL 距离（Kullback-Leibler Divergence）来保证各高阶因子对应的子表征空间相互尽量独立。

### 3.4 注意力机制优化

本工作从注意力机制中获得灵感，将模型优化为如图 3.2 所示的结构。

采用注意力机制来优化该模型是出于两方面的考虑：一方面，通过社会化推荐系统学习得到的用户表征经过长时间的多轮训练，可能会导致邻居之间的用户表征具有比较强的相似性，考虑到在某些推荐场合下可能不希望将注意力更多放在社交影响上，会不利于进行个性化的推荐；另一方面，目前的模型只是实现了用户-物品历史记录上的解耦，并没有实现显式的社交信息上的解耦，而是通过假设表征相似度高则更可能在第  $k$  各因子上互相影响，实际上并没有清晰的解耦步骤。

不过由于本工作中只涉及到两部分表征的加和，并且对于一个推荐系统来说我们认为社交信息的重要程度应该是整体上保持一致的，所以本工作并不涉及到

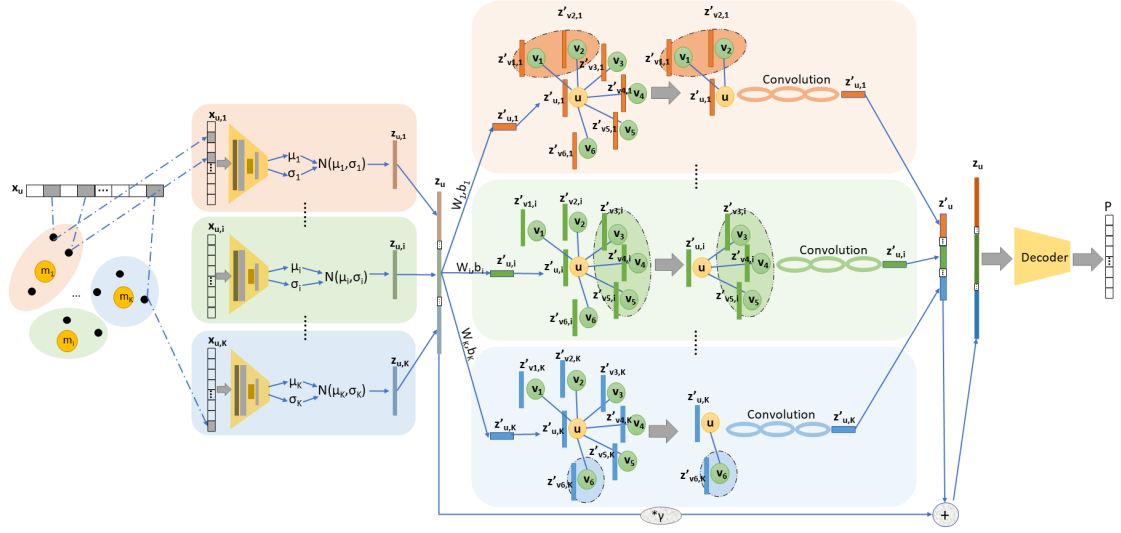


图 3.2 使用注意力机制进行优化后的模型整体结构。

权重计算过程，而只是用一个超参数  $\gamma$  来实现注意力的转移。

通过对比图 3.1 和图 3.2 的结构，可以看到优化后的模型只是在编码模块的路由传播部分与原先不同，增加了社交信息的解耦并将不带社交信息的子表征重新与学习了社交信息的子表征按权重相加。同时，由于图 3.2 中灰色框的部分其实输入输出在含义上是保持一致的，因此多个灰色块完全可以级联相接，所以实际上该路由传播部分可以根据需要调整层数，在一轮训练中进行多次路由传播。

### 3.4.1 社交信息解耦

在学习用户-物品历史记录提取出的表征的基础上再次进行社交信息的解耦。社交信息的解耦需要为社交信息寻找新的子表征空间中心，这里我们参考的是 DisenGCN<sup>[2]</sup> 中进行表征解耦的方法：利用  $d * d$  维的矩阵  $W_k$  以及一个  $d$  维的偏差  $b_k$  来描述社交信息对应于某个因子的子表征空间的中心，通过训练学习  $K$  个不同的矩阵和偏差来实现不同高阶因子对应子空间之间的解耦。所以在将用户表征分离到各因子对应的子空间时的公式如下：

$$z'_{u,k} = \sigma(W_k^T z_u + b_k)$$

$$z'_{u,k} = z'_{u,k} / \|z'_{u,k}\|_2$$

$z'_{u,k}$  即为用户  $u$  在社交信息解耦得到的因子  $k$  上的解耦表征。

### 3.4.2 注意力机制

我们在这里认为整体推荐系统各高阶因子的子表征所构成的子空间的中心可以由从用户-物品历史记录中学习到的用户子表征的子空间中心，以及该表征经过在社交信息上的重新解耦得到的子空间中心的线性组合得到，并且训练中模型可以逐渐适应这一变化。

注意力机制在本模型中的工作机制非常简单：

$$\begin{aligned} z_{u,k} &= \gamma * z_{u,k} + z'_{u,k} \\ z_{u,k} &= z_{u,k} / \|z_{u,k}\|_2 \end{aligned}$$

在这里，我们通过超参数  $\gamma$  来实现注意力在两个不同解耦空间之间的迁移。

## 第 4 章 实验结果

### 4.1 数据集

在本工作实验中使用到的数据集包括两个：Douban<sup>[14]</sup>和 LastFM<sup>[15]</sup>。

Douban 的数据源自于豆瓣电影，豆瓣是一个在中国受到广泛欢迎的社交网站。我使用的 Douban 数据集包含 2848 个用户和 39586 部电影，总共有 894887 个打分数据，平均每个用户对大约 314 部电影打过分。由于本工作采用的是单值的数据，所以我们认为用户对打过分的电影比较感兴趣，将用户打过分的电影对应的特征位置置为 1。Douban 数据集也包含这 2848 个用户在豆瓣上的社交关系，构成了一个社交网络，该社交网络共包含 17885 条边。

LastFM 的数据源于 Last.fm 在线音乐平台，包含 1892 个用户和 17632 个音乐艺术家（包括歌手、作曲家、乐队等等），共 92834 条打分数据，平均每个用户给大约 49 位艺术家打过分。该数据集中包含的对应于这 1892 个用户的社交网络共 12717 条无向边。除此以外，该数据集还提供了一些其他信息，但是在本工作中没有用到。

### 4.2 实现细节

在进行实验时一些超参数设置如下：

表 4.1 实验参数

超参数符号	超参数含义	取值
lr	learning rate, 学习率	1e-3
keep1	keep probability, 编码块中编码部分 dropout 层的激活概率	0.5
keep2	keep probability, 编码块中路由传播部分最后添加的 dropout 层的激活概率	0.65
$\beta$	超参数, 出现在目标函数中	0.2
$\tau$	超参数, 控制 softmax、sigmoid 的输入范围	0.1
$\sigma_0$	超参数, 假设的整体表征空间服从的先验正态分布的方差	0.075



表 4.1 实验参数（续）

超参数符号	超参数含义	取值
$\gamma$	超参数，出现在注意力机制优化部分中	8
kfac	超参数，高阶因子的数量	LastFM: 7, Douban: 9
dfac	超参数，对应于每个高阶因子的子表征的维度	LastFM: 100, Douban: 150
numLayer	在注意力机制优化的模型中，灰色块的级联层数	1
T	路由传播部分的更新轮数	6
nbsz	路由传播时选择的邻居个数	20
early	early stop 的轮数	20
split	从训练数据中分离出的验证集所占比例	0.1

实验中，我们使用 Adam 优化器<sup>[16]</sup>来进行训练。同时，由于这两个数据集中不包含时间戳的信息，所以我们直接将数据集中的部分打分数据抽取出 10% 作为测试集，做为对用户未来选择偏好的预测。

### 4.3 实验结果

本部分将设计的未经优化的社会化推荐系统和用注意力机制优化的社会化推荐系统与 Ma et al.<sup>[1]</sup>在数据集上相对比，所使用的指标包括：

- NDCG@100:

NDCG 值是在推荐系统研究常用的一个衡量推荐排序顺序准确性的指标，其计算方式如下：

$$\begin{aligned}
 DCG@K &= \sum_{i=1}^K \frac{r(i)}{\log_2(i+1)} \\
 IDCG_u &= \sum_{i=1}^K \frac{1}{\log_2(i+1)} \\
 NDCG_u@K &= \frac{DCG_u@K}{IDCG_u} \\
 NDCG@K &= \frac{1}{N} \sum_{u=1}^N NDCG_u@K
 \end{aligned}$$

其中  $r(i)$  是一个权重函数，在本工作的应用场景下，我们认为当我们所得到的推荐列表中的第  $i$  个物品出现在测试集中，则  $r(i) = 1$ ，否则  $r(i) = 0$ 。  $N$  为用户个数。

- Recall@20 及 Recall@50:

Recall 召回率是在机器学习中常用的一个指标，在本实验场景下的语义大致为：在最后推荐列表里 top@K 对应的物品有多少出现在了测试集里。

实验结果如下所示：

表 4.2 实验结果

数据集 (方法)	NDCG@100	Recall@20	Recall@50
Douban(Ma et al. <sup>[1]</sup> )	0.13714	0.10958	0.15044
Douban(初步模型)	0.07915	0.06084	0.08021
Douban(优化模型)	0.14065	0.11292	0.15502
LastFM(Ma et al. <sup>[1]</sup> )	0.21094	0.21517	0.32130
LastFM(初步模型)	0.20061	0.19801	0.30382
LastFM(优化模型)	0.22531	0.23077	0.34227

以上实验数据均由五轮实验结果取平均值得到。

可以看到优化后的模型在三个指标上都显著超过了 Ma et al.<sup>[1]</sup>的结果，并且相较于优化前的模型有非常显著的提高，而优化前的模型实验结果还不如 Ma et al.<sup>[1]</sup>。因此，一方面证实了社交信息可以提升推荐效果；另一方面也说明平衡社交关系和本身用户表征之间的注意力也是很重要的，否则可能会导致用户和邻居之间逐渐趋同，影响推荐效果。

## 第 5 章 结论

本工作研究了如何利用社交网络和解耦机器学习来提升推荐系统可解释性、提高推荐的准确率。为了能够将推荐系统解耦并有效利用社交信息提升推荐准确率，本工作提出了一套解耦的社会化推荐系统结构并用注意力机制对其进行了优化。通过实验表明，经过优化后的社会化推荐系统能够实现推荐效率上的提升并做到将表征解耦为对应于  $k$  个高阶因子的子表征。

然而本工作还有很多改进的空间。本工作所实现的解耦模型的推荐效果与当今前沿的非解耦模型的推荐效果还有一定差距，可能还需要考虑一些其他优化方法来提升模型的推荐效果。同时本工作目前考虑的数据类型均为单值数据且没有考虑一些数据上额外的标签信息，未来可以对于如何设计可以处理带权值的社交网络、分数非单值的历史记录的解耦社会化推荐系统开展研究，并在设计的过程中考虑如何加入对于额外标签的学习。

## 插图索引

图 2.1	Ma et al. <sup>[1]</sup> 的模型结构。原图来自论文 <sup>[1]</sup> 。 .....	9
图 2.2	DisenGCN 的模型结构。原图来自论文 <sup>[2]</sup> 。 .....	10
图 3.1	初步模型的整体结构。 .....	11
图 3.2	使用注意力机制进行优化后的模型整体结构。 .....	15

## 表格索引

表 4.1	实验参数 .....	17
表 4.2	实验结果 .....	19

## 参考文献

- [1] MA J, ZHOU C, CUI P, et al. Learning disentangled representations for recommendation[J/OL]. CoRR, 2019, abs/1910.14238. <http://arxiv.org/abs/1910.14238>.
- [2] MA J, CUI P, KUANG K, et al. Disentangled graph convolutional networks[C/OL]// CHAUDHURI K, SALAKHUTDINOV R. Proceedings of Machine Learning Research: volume 97 Proceedings of the 36th International Conference on Machine Learning. PMLR, 2019: 4212-4221. <https://proceedings.mlr.press/v97/ma19a.html>.
- [3] BAHDANAU D, CHO K, BENGIO Y. Neural machine translation by jointly learning to align and translate[M/OL]. arXiv, 2014. <https://arxiv.org/abs/1409.0473>. DOI: 10.48550/ARXIV.1409.0473.
- [4] WANG X, HE X, WANG M, et al. Neural graph collaborative filtering[J/OL]. CoRR, 2019, abs/1905.08108. <http://arxiv.org/abs/1905.08108>.
- [5] HE X, DENG K, WANG X, et al. Lightgcn: Simplifying and powering graph convolution network for recommendation[J/OL]. CoRR, 2020, abs/2002.02126. <https://arxiv.org/abs/2002.02126>.
- [6] YU J, YIN H, XIA X, et al. Graph augmentation-free contrastive learning for recommendation [J/OL]. CoRR, 2021, abs/2112.08679. <https://arxiv.org/abs/2112.08679>.
- [7] WU J, WANG X, FENG F, et al. Self-supervised graph learning for recommendation[M/OL]. New York, NY, USA: Association for Computing Machinery, 2021: 726–735. <https://doi.org/10.1145/3404835.3462862>.
- [8] YU J, YIN H, GAO M, et al. Socially-aware self-supervised tri-training for recommendation [J/OL]. CoRR, 2021, abs/2106.03569. <https://arxiv.org/abs/2106.03569>.
- [9] YU J, YIN H, LI J, et al. Self-supervised multi-channel hypergraph convolutional network for social recommendation[J/OL]. CoRR, 2021, abs/2101.06448. <https://arxiv.org/abs/2101.06448>.
- [10] WU L, SUN P, FU Y, et al. A neural influence diffusion model for social recommendation [J/OL]. CoRR, 2019, abs/1904.10322. <http://arxiv.org/abs/1904.10322>.
- [11] ZHU Y, XU Y, YU F, et al. Disentangled self-attentive neural networks for click-through rate prediction[J/OL]. CoRR, 2021, abs/2101.03654. <https://arxiv.org/abs/2101.03654>.
- [12] WANG X, JIN H, ZHANG A, et al. Disentangled graph collaborative filtering[J/OL]. CoRR, 2020, abs/2007.01764. <https://arxiv.org/abs/2007.01764>.

- [13] CUI Z, YU F, WU S, et al. Disentangled item representation for recommender systems[J/OL]. CoRR, 2020, abs/2008.07178. <https://arxiv.org/abs/2008.07178>.
- [14] ZHAO G, QIAN X, XIE X. User-service rating prediction by exploring social users' rating behaviors[J/OL]. IEEE Transactions on Multimedia, 2016, 18(3): 496-506. DOI: 10.1109/TMM.2016.2515362.
- [15] CANTADOR I, BRUSILOVSKY P, KUFLIK T. Second workshop on information heterogeneity and fusion in recommender systems (hetrec2011)[C/OL]//MOBASHER B, BURKE R D, JANNACH D, et al. Proceedings of the 2011 ACM Conference on Recommender Systems, RecSys 2011, Chicago, IL, USA, October 23-27, 2011. ACM, 2011: 387-388. <https://doi.org/10.1145/2043932.2044016>.
- [16] KINGMA D P, BA J. Adam: A method for stochastic optimization[C/OL]//BENGIO Y, LECUN Y. 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings. 2015. <http://arxiv.org/abs/1412.6980>.