

4 Large Language Model Augmented Narrative Driven Recommendations

4.1 论文来源

Mysore S, McCallum A, Zamani H. Large language model augmented narrative driven recommendations[C]//Proceedings of the 17th ACM Conference on Recommender Systems. 2023: 777-783.

4.2 论文代码

<https://github.com/iesl/narrative-driven-rec-mint/>

代码不完整，不支持复现。

4.3 论文笔记

4.3.1 模型总结

叙事驱动的推荐（NDR）提出了一个信息访问问题，其中用户通过对其偏好和背景的详细描述来征求推荐，例如，旅行者在描述他们的喜好和旅行环境的同时征求兴趣点的推荐。随着基于自然语言的搜索和推荐系统对话界面的兴起，这些需求变得越来越重要。然而，NDR 缺乏丰富的模型训练数据，而当前的平台通常不支持这些请求。幸运的是，经典的用户-项目交互数据集包含丰富的文本数据，例如评论，它们通常描述用户偏好和上下文 - 这可以用于引导 NDR 模型的训练。在这项工作中，我们探索使用大型语言模型（LLM）进行数据增强来训练 NDR 模型。我们使用大语言模型通过小样本提示从用户-项目交互中编写综合叙述查询，并在综合查询和用户-项目交互数据上训练 NDR 检索模型。我们的实验表明，这是训练小参数检索模型的有效策略，该模型在叙事驱动推荐方面优于其他检索和大模型基线。

4.3.2 背景介绍

针对用户的个性化推荐系统是多个行业规模平台的重要组成部分[16,17,46]。这些系统的功能是根据用户之前在平台上的交互推断用户的兴趣，并根据这些推断的兴趣做出推荐。虽然基于历史交互的推荐是有效的，但用户获得推荐通常从对他们想要的目标项目的模糊想法开始，或者可能需要根据使用上下文的推荐，而历史交互数据中经常缺失（图 1）。



Looking for restaurants and food in Rotterdam, NL
Hi all - my father and I will be visiting your fair city next week, and we're looking for a place to have dinner on a Sunday night. Are there any places that you would recommend? We like tasty food, good beer, and a non-stuffy atmosphere. Considerations: Open on Sunday (I gather this is somewhat limiting?) Our hotel is right near the Leuvehaven metro stop, and walking distance would be nice Ideally not Asian(-inspired) food - we're from southeast Asia and would rather try something different. Totally up for everything else though! Thank you!.

图 1：征求兴趣点推荐的叙述性查询示例。该查询描述了用户的偏好及其请求的上下文。

Figure 1: An example narrative query soliciting point of interest recommendations. The query describes the users preferences and the context of their request.

在这些场景中，用户通常通过描述其广泛兴趣和背景的长篇叙述性查询来征求建议。此类信息访问任务已被研究为从书籍 [5] 和电影 [18] 到兴趣点 [1] 等项目的叙事驱动推荐 (NDR)。Bogers 和 Koolen [5] 指出这些叙述请求在讨论论坛和几个 Reddit 子版块中很常见¹，但是，当前的推荐系统缺乏对这些复杂的自然语言查询的支持。

然而，随着信息访问任务对话界面的出现，对复杂 NDR 任务的支持可能变得必要。在这种情况下，最近的工作注意到与更传统的搜索界面相比，复杂和主观的自然语言请求有所增加 [13, 34]。此外，具有强大语言理解能力的大型语言模型 (LLM) 的出现提供了满足此类复杂请求的潜力 [9, 33]。这项工作探索了重新利用历史用户项目推荐数据集的潜力，这些数据集传统上用于训练协同过滤推荐器，并使用 LLM 来支持 NDR。

具体来说，给定用户的交互 D_u ，项目及其随附文本文档（例如评论、描述）的交互 $D_u = \{d_i\}_{i=1}^{N_u}$ ，从用户-项目交互数据集 I 中选择，我们提示 InstructGPT，一个 175B 参数的 LLM，创作基于 D_u 的综合叙述查询 q_u （图 2）。

3 few-shot examples:

A user likes these recommendations: Mission BBQ in Deptford Township, Keswick Tavern in Glenside ...

The user wrote these reviews: This place is always packed! Keswick Tavern is a go to if you're looking for great food and large alcohol/beer selection ...

In response to the request on Reddit: Hi Philadelphia area friends, I will be moving Philadelphia soon and I am looking for some popular local restaurants I can turn into my staples ...

Target completion for LLM:

A user likes the following recommendations: Rue De La Course in New Orleans, Swirl Wine Bar & Market in New Orleans ...

The user wrote the following reviews: The wide variety of tea is impressive. On Fridays they offer free wine tastings ...

In response to the request on Reddit: LLM completion ...

图 2: Mint 中使用的提示格式, 用于根据用户-项目与大型语言模型的交互生成综合叙述查询。

Figure 2: The format of the prompt used in MINT for generating synthetic narrative queries from user-item interaction with a large language model.

由于我们预计查询 q_u 是有噪声的, 并且不能完全代表所有用户评论, 因此根据给定用户文档 d_i 的语言模型分配的 q_u 可能性, 对 D_u 进行过滤以仅保留一小部分评论。然后, 对基于预训练 LM 的检索模型 (110M 参数) 进行微调, 以用于综合查询和过滤评论的检索。

我们的方法 (我们称之为 Mint2) 是根据观察得出的, 虽然叙述性查询和建议经常在在线讨论论坛中提出, 并且可以作为培训数据, 但这些帖子的数量以及它们可用的领域的多样性明显小于被动收集的用户-项目交互数据集的大小和多样性。例如, 虽然 Bogers 和 Koolen [5] 注意到 LibraryThing 讨论论坛上有近 25,000 个对图书的叙述请求, 但 Goodreads 的公开用户项目交互数据集包含 46 万用户与近 220 万本书的交互 [43]。

我们在一个公开的兴趣点推荐测试集中对 Mint 进行了实证评估: pointrec [1]。为了训练我们的 NDR 模型, 我们根据 Yelp 的用户-项目交互数据集生成综合训练数据。使用 Mint 训练的模型 (110M 参数) 显著优于多个基线模型, 并与自回归生成推荐的更大的 LLM 基线的性能相匹配。代码和合成数据集可用。

4.3.3 方法

我们提出的 NDR 方法 Mint, 通过使用 LLM 作为查询生成模型, 将大量可用的用户-项目交互数据集 $I = \{(u, \{d_i\}_{i=1}^{N_u})\}$ 重新调整为检索模型的训练数据编写叙述性查询 $q_u : D = \{(q_u, \{d_i\}_{i=1}^{N_u})\}$ 。然后, 在合成数据集 D 上训练检索模型 (图 3)。

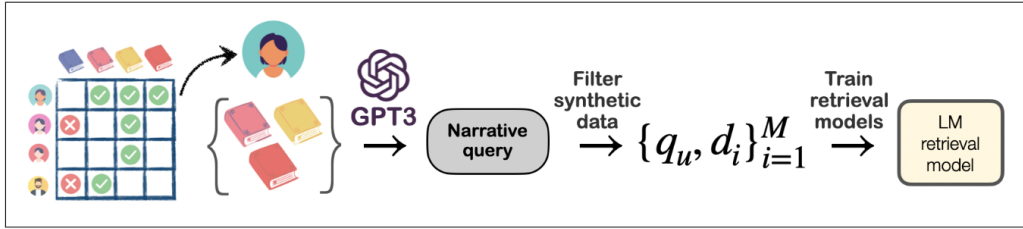


Figure 3: MINT re-purposes readily available user-item interaction datasets commonly used to train collaborative filtering models for narrative-driven recommendation. This is done by authoring narrative queries for sets of items liked by a user with a large language model. The data is filtered with a smaller language model and retrieval models are trained on the synthetic queries and user items.

图 3: Mint 重新利用了现成的用户-项目交互数据集, 这些数据集通常用于训练协作过滤模型以实现叙事驱动的推荐。这是通过使用大型语言模型为用户喜欢的项目集编写叙述性查询来完成的。使用较小的语言模型过滤数据, 并根据合成查询和用户项目训练检索模型。

4.3.3.1 LLM的叙述性查询

为了在 I 中为用户编写叙述性查询 q_u , 我们使用 175B 参数 InstructGPT 模型作为查询生成模型 QGen。我们在 QGen 的提示中包含交互项的文本 $\{d_i\}_{i=1}^{N_u}$, 并指示它编写叙述性查询 (图 2)。为了提高生成的查询的一致性并获得正确格式的输, 我们根据 3 个主题不同的用户的交互项手动编写叙述性查询, 并将其包含在 QGen 的提示中。整个数据集 I 使用相同的三个小样本示例, 并且从 I 中选择三个用户。基于用户交互生成叙述性查询也可以被视为用于生成自然语言用户配置文件的多文档摘要的一种形式[36]。

4.3.3.2 过滤综合查询的项目

由于我们期望用户项目捕获其兴趣的多个方面, 并且生成的查询仅捕获这些兴趣的子集, 因此在将其用于训练检索模型之前, 我们仅保留 $\{d_i\}_{i=1}^{N_u}$ 中存在的一些项目。为此, 我们使用预先训练的语言模型来计算给定每个用户项目 PLM ($q_u | d_i$) 的查询的可能性, 并且只保留 q_u 的前 M 个高评分项目 (即只提及用户感兴趣的 top k item), 这会导致对于我们的 NDR 检索模型: $\{(q_u, d_i)_{i=1}^M\}$, 每个用户有 M 个训练样本。在我们的实验中, 我们使用具有 3B 参数的 FlanT5 [14] 进行计算, 并遵循 Sachan 等人的方法计算 PLM ($q_u | d_i$)。请注意, 我们对 PLM ($q_u | d_i$) 的使用代表了一种通常用于即席搜索的查询似然模型, 并且最近被证明在与大型预训练语言模型一起使用时是一种有效的无监督重新排序方法 [40]。

4.3.3.3 训练检索模型

我们在生成的合成数据集上训练 NDR 的双编码器和交叉编码器模型 - 搜索任务中常用的模型。双编码器通常用作大量项目集中的可扩展第一阶段排名器。另一方面, 交叉编码器允许查询和项目之间更丰富的交互, 并用作第二阶段的重排序模型。对于这两个模型, 我们都使用具有 110M 参数的预训练 Transformer 语言模型架构, MPnet, 一个类似于 Bert [41] 的模型。

4.3.4 实验和结果

表 1 列出了所提出的方法与基线相比的性能。此处, 粗体数字表示表现最佳的模型, 上标表示使用 $p < 0.05$ 的双边 t 检验计算出的统计显著性。

Table 1: Performance of the proposed method, MINT, for point-of-interest recommendation on POINTREC. The superscripts denote statistically significant improvements compared to specific baseline models.

表 1: 所提出的方法 Mint 在 Pointrec 上的兴趣点推荐的性能。上标表示与特定基线模型相比在统计上显著的改进。

Model	Parameters	POINTREC					
		NDCG@5	NDCG@10	MAP	MRR	Recall@100	Recall@200
¹ BM25	-	0.2682	0.2464	0.1182	0.2685	0.4194	0.5429
² Contriver	110M	0.2924	0.2776	0.1660	0.3355	0.4455	0.5552
³ MPNet-1B	110M	0.3038	0.2842	0.1621	0.3566	0.4439	0.5657
⁴ BERT-MSM	110M	0.3117	0.2886	0.1528	0.3320	0.4679	0.5816
⁵ Grounded LLM	175B+110M	0.3558	0.3251	0.1808	0.3861	0.4797	0.5797
⁶ UPR	110M+3B	0.3586	0.3242	0.1712	0.4013	0.4489	0.5552
BiEnc-MINT	110M	0.3489 ¹	0.3263 ¹	0.1890 ¹	0.3982 ¹	0.4914 ¹	0.6221
CrEnc-MINT	2×110M	0.3725¹²	0.3489¹²	0.2192¹⁴	0.4317¹	0.5448¹²³	0.6221

4.4 启发

在本文中，提出了 Mint，一种用于叙事驱动推荐（NDR）任务的数据增强方法。Mint 通过使用 175B 参数大语言模型来编写长篇叙述性查询，同时根据用户喜欢的项目文本进行调节，重新利用 NDR 的历史用户-项目交互数据集。这种利用大模型进行数据处理的方式十分新颖，可以考虑在未来的工作中使用类似的操作来得到高质量的数据集。例如在之前看的考虑因果的推荐系统中，使用自监督学习的方法来进行序列间item的因果学习，我们可以使用大模型实现对item序列之间因果关系的判断，进而将大模型的判断结果作为一个特征输入到下游的推荐系统模型中完成序列推荐。

5 Beyond Co-occurrence: Multi-modal Session-based Recommendation

5.1 论文来源

Zhang X, Xu B, Ma F, et al. Beyond co-occurrence: Multi-modal session-based recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2023.

5.2 论文代码

<https://github.com/Zhang-xiaokun/MMSBR>

5.3 论文笔记

基于会话的推荐致力于基于短会话来表征匿名用户的偏好。现有的方法主要集中于挖掘会话内项目 ID 所暴露的有限项目共现模式，而忽略了吸引用户与某些项目进行交互的页面上显示的丰富的多模态信息。一般来说，多模态信息可以分为两类：描述性信息（例如物品图像和描述文本）和数字信息（例如价格）。在本文中，我们的目标是通过对上述多模态信息进行整体建模来改进基于会话的推荐。从多模态信息中揭示用户意图主要存在三个问题：（1）如何从具有不同噪声的异构描述信息中提取相关语义？（2）如何融合这些异构的描述信息来综合推断用户兴趣？（3）如何处理数值信息对用户行为的概率影响？为了解决上述问题，我们提出了一种新颖的基于多模式会话的推荐（MMSBR），它在统一的框架下对描述性信息和数字信息进行建模。具体来说，设计了伪模态对比学习来增强描述信息的表示学习。然后，提出了分层枢轴变压器来融合异构描述信息。此外，我们用高斯分布表示数值信息，并设计 Wasserstein 自注意力来处理概率影响模式。对三个真实世界数据集的广泛实验证明了所提出的 MMSBR 的有效性。进一步的分析也证明我们的 MMSBR 可以有效缓解 SBR 的冷启动问题。

直观上，页面上显示的多模态信息驱使用户与某些项目进行互动。如图1所示，用户通常在查看商品图像、阅读描述文本并检查价格后做出决定。由于依赖不同的载体来传达特定的项目特征，多模态信息可以分为两类:描述性信息和数值信息。



Fig. 1: A user makes the decision after evaluating all multi-modal information displayed on pages including item images, description text and price.

描述性信息是用图像和文字来描绘物品，可以直观地描述物品的一些特征，如款式、颜色、材质等。对于数字信息，即价格，它通过实数传递物品的抽象价值。在大多数情况下，如图 1 所示，用户不会单击某个项目，除非她对项目的各个方面都感到满意。显然，上述多模态信息共同决定了用户的选择。

本文提出了一种新颖的基于多模式会话的推荐（MMSBR），它可以定制确定性和概率模型来分别处理描述性和数字信息。在确定性建模中，我们设计了一种伪模态对比学习来细化描述性信息表示。特别是，对比学习通过将语义相似（正）对拉近，同时将不相似（负）对拉开来增强表示学习[15]。

由于项目的不同模态引用相似的内容，因此直观地将它们视为正对(对比学习中的正+对)来解决噪声问题。然而，不同模态之间存在语义差距，因此不适合直接对比它们。为了解决这个问题，我们建议利用一种模态通过数据生成技术以另一种模态生成伪信息（即伪模态）。然后，在同一语义空间中对齐的实际模态和伪模态在对比学习中用作正对，以减轻图像和文本中存在的噪声。

此外，为了融合描述性信息，我们在确定性建模中提出了分层枢轴变换器Transformer。凭借对序列中复杂关系进行建模的能力，Transformer 结构已被证明可以有效地合并多模态信号 [16]、[17]。受此启发，我们进一步在每个变压器层中创建一个枢轴，作为有价值信息的混合器，以管理异构信息的融合。Pivot 在 Transformer 操作下分层提取并整合来自图像和文本的有用信息。然后，我们将枢轴视为描述性信息的全面嵌入。

在概率建模中，我们首先将商品价格表示为高斯分布嵌入，这使得 MMSBR 能够感知商品价格的范围属性。然后开发 Wasserstein 自注意力来处理价格分布嵌入，以获得用户可接受的价格范围。由于能够区分高斯分布之间的差异，Wasserstein 距离 [18]、[19] 被用于 Wasserstein 自注意力，以确定价格分布嵌入之间的相关性。最后，所提出的 MMSBR 通过评估页面上显示的整个多模式信息为用户提供个性化服务。

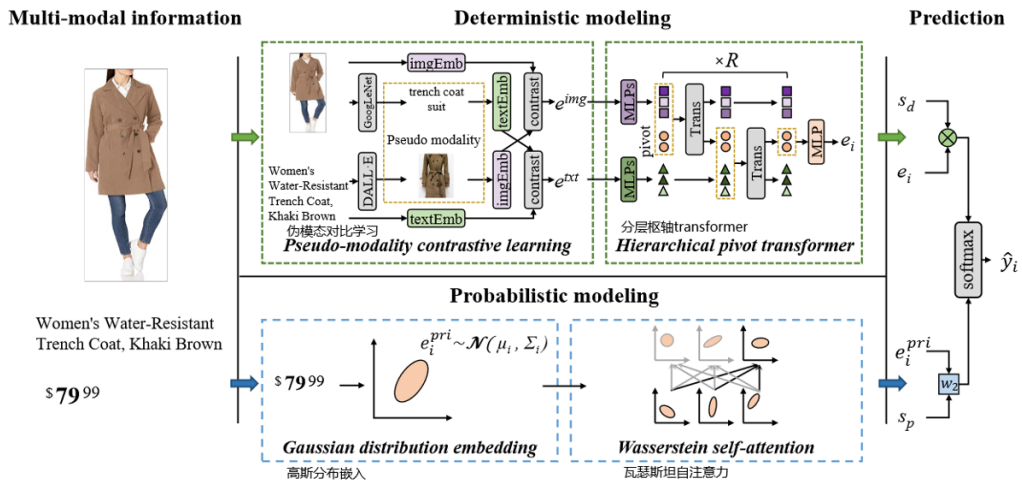

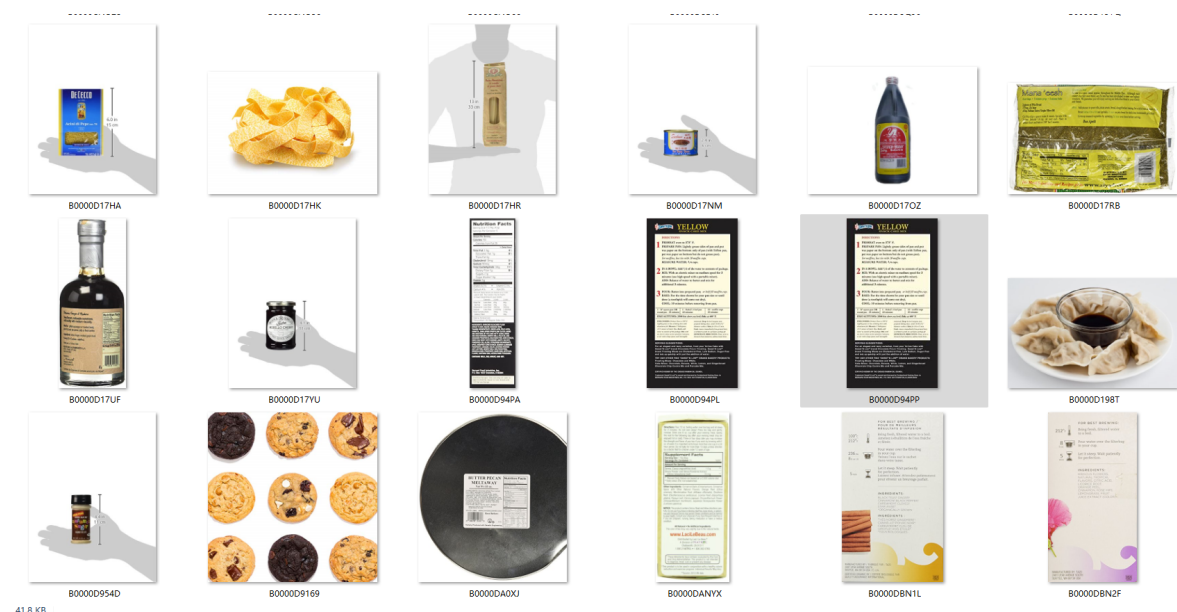


Fig. 2: The proposed MMSBR customizes deterministic and probabilistic modeling to handle descriptive and numerical

5.4 启发

在推荐的过程中使用图片信息，可能需要图片中的商品尽可能明确，如果像是框架图中的照片，在由的伪模态对比学习中，可能会引入噪声，因为图片中不仅仅有上衣，还有裤子。我觉得多模态推荐的难点可能是对数据集的高要求。对于本文中所采用的数据集的图片模态，我观察了一下发现，亚马逊中item的图片质量良莠不齐，有的确实可以很好的描述商品，但有的可能仅仅是一个商品的说明书，推荐模型可能完全没有办法从说明书图片中获得关于item的细节信息。



6 Bi-preference Learning Heterogeneous Hypergraph Networks for Session-based Recommendation

6.1 论文来源

Zhang X, Xu B, Ma F, et al. Bi-Preference Learning Heterogeneous Hypergraph Networks for Session-based Recommendation[J]. ACM Transactions on Information Systems, 2023, 42(3): 1-28.

6.2 论文代码

<https://github.com/Zhang-xiaokun/BiPNet>

代码不完整，缺少数据处理部分的代码。

6.3 论文笔记

基于会话的推荐旨在根据匿名行为序列预测下一个购买的商品。大量经济学研究表明，商品价格是影响用户购买决策的关键因素。不幸的是，现有的基于会话的推荐方法仅旨在捕获用户的兴趣偏好，而忽略了用户的价格偏好。实际上，主要有两个挑战阻碍我们获得价格优惠。首先，价格偏好与各种商品特征（即类别和品牌）高度相关，这要求我们从异构信息中挖掘价格偏好。其次，价格偏好和兴趣偏好是相互依存的，共同决定用户的选择，因此我们需要在意图建模时共同考虑价格偏好和兴趣偏好。为了应对上述挑战，我们提出了一种用于基于会话的推荐的双偏好学习异构超图网络（BiPNet）的新方法。具体来说，设计了具有三级卷积的定制异构超图网络，以从项目的异构特征中捕获用户价格和兴趣偏好。此外，我们开发了一种双偏好学习模式来探索价格和兴趣偏好之间的相互关系，并在多任务学习架构下共同学习这两种偏好。对多个公共数据集的广泛实验证实了 BiPNet 相对于竞争基线的优越性。其他研究也支持价格对于任务至关重要的观点。

用户价格偏好是多样化的，并且在很大程度上取决于商品的某些特征，即类别和品牌[44, 63]。举一个常见的例子，用户可能会购买昂贵的相机用于工作，同时她也可以购买便宜的透明胶带用于日常生活。这个例子表明，用户的价格偏好可能会根据商品类别发生巨大变化。此外，一些用户愿意花额外的钱购买知名品牌的商品，这在经济学中称为品牌效应[7]。这意味着商品品牌也会影响用户的价格偏好。显然，当我们在 SBR 中对用户价格偏好进行建模时，应考虑相关类别和品牌。在这种情况下，需要各种信息来表征用户行为，例如一系列商品、商品价格、商品类别和商品品牌。这些异构信息给现有方法（例如参考文献[4,16,25]）带来了棘手的挑战。因为它们都是为了模拟单一类型信息（即项目 id）而开发的。因此，要将价格纳入 SBR，第一个挑战是如何复制这种异构信息。

使用通用异构图来处理异构信息是直观的[18, 30]。然而，现有的基于异构图的方法只能建模成对（二元）关联[41]，例如（类别，价格）或（品牌，价格）之间的关系。在我们的设置中，不同节点之间存在复杂的高阶关系，例如，<价格、类别、品牌>中的三元关联共同影响商品价格语义。因此，当前基于异构图的方法在我们的设置中是无效的。此外，受过度平滑问题的限制，基于异构图的方法无法对会话中远程节点之间的依赖关系进行建模[4]。过度平滑问题通常是指在图卷积神经网络（Graph Convolutional Networks, GCN）或其他基于图的方法中，对于节点特征的更新或聚合过程中，会导致节点之间的差异逐渐减小，最终使得整个图的表征变得过于平滑，失去了原始节点之间的细节信息。这可能导致模型在学习过程中无法充分捕捉到图结构中的局部模式和变化。

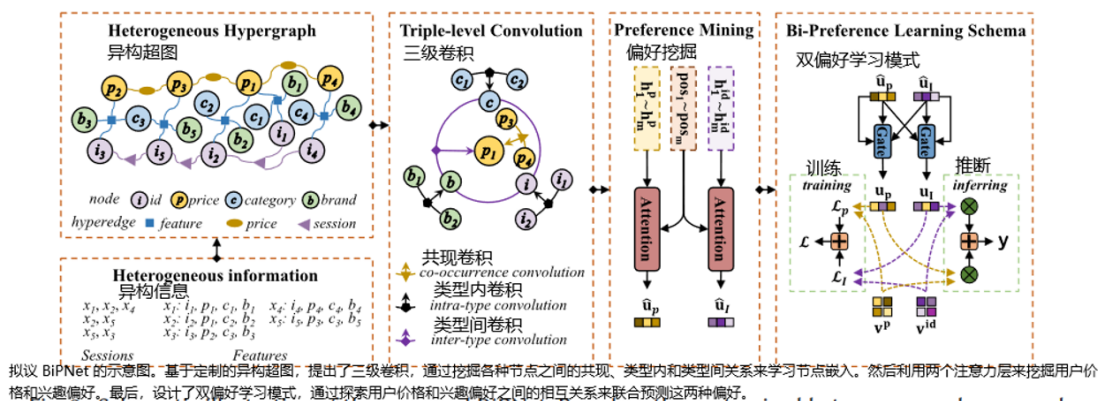
幸运的是，最近设计了超边可以连接多个节点的超图来捕获任意距离的高阶数据关系[54, 64]。利用异构图和超图的优点，我们提出异构超图网络，它能够捕获各个节点之间复杂的高阶依赖关系，并在 SBR 中复制异构信息。

具体来说，我们在异构超图中考虑的异构信息包括与用户价格和兴趣偏好直接相关的商品 ID、商品价格、商品类别和商品品牌。此外，设计了三种类型的超边，即特征超边、价格超边和会话超边，以促进异构超图中的价格和兴趣偏好学习。基于构建的异构超图，我们设计了三级卷积来学习节点嵌入，其中强调节点之间的三种关系，即共现、类型内和类型间关系。之后，通过注意力层获得用户的价格和兴趣偏好。

用户价格偏好和兴趣偏好相互依存，共同决定用户选择。用户在购物时进行价格与利益的权衡是很常见的。此外，经济学家还使用价格弹性来描述用户愿意花在某件商品上的钱随着她对该商品的兴趣而波动的现象[1]。这表明价格和兴趣偏好相互影响，共同影响用户的选择。举个例子，用户有可能因为强烈的兴趣而购买了昂贵的商品，即使该商品的价格比她预期的要高。在这种情况下，用户由于兴趣的影响而改变了她的价格偏好。同样，用户也会因为价格低廉而购买兴趣不大的商品（这正是促销期间经常发生的情况）。显然，为了预测用户购买决策，我们必须探索价格和兴趣偏好之间的相互关系，并对这两种偏好进行联合建模。然而，现有的 SBR 方法[4,16,25,40,52-54]仅使用项目 ID 作为监督信号来建模用户兴趣偏好，而既不能明确捕获用户价格偏好，也不能充分处理价格偏好与用户兴趣偏好之间的丰富依赖关系。兴趣偏好决定用户的选择。因此，我们面临的另一个挑战是如何准确地对这两种偏好下的用户购买意图进行建模。

我们应该探索价格和兴趣偏好之间的相互关系，并将两者考虑在内来预测用户行为。提出多任务学习模式来联合处理多个相关任务，其中可以利用这些任务之间的相互关系来同时提高所有任务的模型能力[62]。受此启发，我们提出了双偏好学习模式。具体来说，除了使用商品 ID 来指导模型学习用户兴趣偏好外，双偏好学习模式还利用目标商品的价格作为另一个监督信号来制定用户价格偏好。在多任务学习的模式下，它可以探索价格和兴趣偏好之间的相互关系并同时预测两者。最终，我们根据学习到的价格偏好和兴趣偏好揭示用户的购买意图。

我们提出了一种新颖的 BiPNet 来改进 SBR。在 BiPNet 中，设计了定制的异构超图网络来处理异构信息以进行价格和兴趣偏好学习。我们还开发了一种双偏好学习模式，通过探索价格和兴趣偏好之间丰富的相互关系来同时预测价格和兴趣偏好。



6.4 启发

我看过亚马逊的数据集，就每个item来说，除了itemID，还有很多其他的属性，我一直关注的问题是如何能够利用这些item的属性，提高推荐系统的精度。将价格作为推荐系统的一个输入，或是多模态推荐将图片作为推荐系统的一个输入，都是为了能够帮助推荐系统从与用户的交互中获得更多的有效信息，精准建模用户的各种偏好，从而达到提高推荐精度的目的。在本文中，定义了异构图来使得item的类型和价格信息可以和itemID一起被输入到模型中，但在经过三层卷积等一系列从异构图中获取信息的模块之后，user-item交互中的时序信息其实是被淡化了，如何能够在有效捕获时序特征的同时考虑到item的多种属性，是一个还没有解决的问题。

关于大模型，微调的话，现在推荐的数据集其实是有点不足的，我发现的论文是直接使用大模型作为推荐系统或者embedding层，目前也没有很好的想法，可以等等看今年新的顶会论文看看大家的工作。