arXiv:2302.03497v2 [cs.IR] 8 Feb 2024

MMRec：简化多模态推荐

Xin Zhou xin.zhou@ntu.edu.sg 阿里巴巴—南洋理工大学新加坡联合研究院 南洋理工大学 新加坡，新加坡

摘要

本文介绍了一个用于多模态推荐的开源工具箱 MMRec。MMRec 简化并规范了多模态推荐模型的实现与比较流程。MMRec 旨在提供一个统一且可配置的平台，最大限度地减少实现和测试多模态推荐模型所需的工作。它支持多模态模型融合多种模态的信息，涵盖从传统矩阵分解到现代基于图的算法。相关文档、示例和源代码可在 https://github.com/enoche/MMRec 获取。

CCS 概念

• **信息系统** → **推荐系统**；**多媒体与多模态检索**。

关键词

多模态推荐，基准。

ACM 参考格式：

Xin Zhou. 2023. MMRec：简化多模态推荐。发表于*ACM*亚洲多媒体研讨会（*MMAsia '23* 研讨会），*2023*年*12*月*6*日至*8*日，台湾台南。ACM，纽约，纽约州，美国，2页。https://doi.org/10.1145/ 3611380.3628561

1 引言

多模态推荐模型正成为当前研究社区的研究趋势，原因如下：

- 多模态信息的普及（例如图像、文本和视频）；
- 在推荐中利用多模态信息可以缓解交互数据的稀疏性 [6–8, 11]；
- 自然语言处理和计算机视觉领域近期在多模态学习方面的研究日趋成熟 [1, 3]。

因此，多模态推荐范式逐渐演变为数字媒体平台不可或缺的基石。这一演进使这些平台能够向用户提供量身定制的推荐。为此，平台会同时分析历史用户交互以及物品的多种模态 [6]。与仅利用用户-物品交互进行推荐的传统推荐系统 [2, 5, 9, 10] 相比，多模态模型涉及多个流程：预处理来自多种模态的信息（例如，*k*-评分用户/物品筛选、数据划分、多模态信息向量化，以及物品 ID 与其多模态信息对齐）、融合多模态信息等。一方面，这些繁琐的流程阻碍了多模态推荐研究的进展。另一方面，预处理方法种类繁多，使得模型的性能难以复现，也难以与其他模型进行公平比较。

为解决这些问题，我们推出了 MMRec，这是一款开源工具箱，旨在简化多模态推荐研究。MMRec 提供了一套完整的工具箱，涵盖数据预处理、多模态推荐模型、多模态信息融合和性能评估，以降低实现新模型或基线方法并对其进行比较的成本。MMRec 的目标是建立一个基准测试系统，确保以高效省力的方式公平比较多模态模型。该工具箱具有很强的可配置性，且易于使用。

2 架构与关键特性

图 1 展示了 MMRec 的模块。需要指出的是，MMRec 的输入包括多模态原始数据和用户-物品交互文件。

数据封装。 MMRec 首先对原始数据进行预处理，并将用户交互和多模态信息封装到 PyTorch 的 DataLoader 中。原始数据的格式与 Amazon Review Data [1] 一致。MMRec 执行 *k*-核过滤，以保留至少有 *k* 次交互的用户和物品，并将多模态信息与保留的物品对齐。随后，它将全部交互划分为训练集/验证集/测试集。借助预训练的多模态模型（例如 Transformer），将多模态信息的原始特征向量化为数值。

训练器。 MMRec 提供多种优化器来训练模型。MMRec 支持单模态和多模态模型。在 MMRec 中，可以轻松融合来自各个模态的信息，用于多模态推荐。在当前版本中，MMRec 支持四种常见模态：文本、图像、音频和视频。MMRec 为所有模型统一了训练接口。自定义模型只需实现两个函数：

计算_损失： 模型的核心部分，用于定义如何根据模型图流程生成损失。**完整_排序_预测：** 此函数用于预测用户对物品的排序。

评估。 此模块提供一系列推荐系统中常用的指标，例如召回率、归一化折损累计增益、平均准确率等。,。

[1] http://jmcauley.ucsd.edu/data/amazon/links.html

本作品采用知识共享署名 4.0 国际许可协议授权。

MMAsia '23 研讨会，*2023*年*12*月*6*日至*8*日，台湾台南 © 2023 版权所有归所有者/作者所有。ACM ISBN 979-8-4007-0326-3/23/12。
https://doi.org/10.1145/3611380.3628561

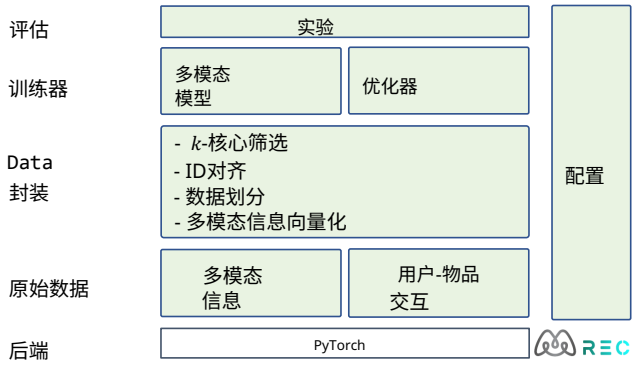

图 1：MMRec 的架构。MMRec 由 4 个模块组成，涵盖从原始数据预处理到模型性能评估的各个环节。为确保模型的可复现性和性能一致性，MMRec 以原始数据为输入。

表 1：多模态推荐框架的比较。

框架	支持的功能				多模态融合
	Text	图像	音频	视频	
Cornac	✓	✓	✗	✗	✗
MMRec	✓	✓	✓	✓	✓

值得注意的是，所有模块都可以通过修改配置文件进行自定义和配置。所有更改都会在**配置**模块中体现并加载。**配置**模块还支持对模型的超参数进行网格搜索。训练完成后，所有超参数组合的结果都会汇总并报告给最终用户。通过在每种超参数组合中重置 DataLoader 中的随机种子和原始数据，确保模型结果具有可复现性。

3 与相关工作的比较

据我们所知，Cornac [4] 是唯一支持推荐中多模态信息的开源库。尽管我们发现 40+ 算法已在 Cornac 中实现，但其中大多数都是未利用多模态信息的通用推荐系统。此外，Cornac 现阶段在模态融合方面存在局限。具体而言，Cornac 无法整合来自多个模态的多模态信息（即文本、图像、音频、视频）。当前版本的 MMRec 支持 10+多模态推荐模型以及种类广泛的通用协同过滤模型。Cornac 与 MMRec 的详细比较见表 1。

4 结论

我们介绍了 MMRec，这是一个多模态推荐库，旨在帮助研究人员实现其模型，并将结果与最先进的基线方法进行比较。完整文档和源代码可访问 https://github.com/enoche/MMRec。

参考文献

[1] Tadas Baltrušaitis、Chaitanya Ahuja 和 Louis-Philippe Morency。2018。多模态机器学习：综述与分类体系。*IEEE*模式分析与机器智能汇刊41，2（2018），423–443。[2] Jiazheng Jing、Yinan Zhang、Xin Zhou 和 Zhiqi Shen。2023。捕捉流行度趋势：一种用于增强物品推荐的简易非个性化方法。*arXiv*预印本 *arXiv:2308.08799*（2023）。[3] Alec Radford、Jong Wook Kim、Chris Hallacy、Aditya Ramesh、Gabriel Goh、Sandhini Agarwal、Girish Sastry、Amanda Askell、Pamela Mishkin、Jack Clark 等。2021。通过自然语言监督学习可迁移视觉模型。载于国际机器学习大会。PMLR，8748–8763。[4] Aghiles Salah、Quoc-Tuan Truong 和 Hady W Lauw。2020。Cornac：多模态推荐系统的比较框架。机器学习研究杂志21，95（2020），1–5。[5] Lingzi Zhang、Yong Liu、Xin Zhou、Chunyan Miao、Guoxin Wang 和 Haihong Tang。2022。基于扩散的图对比学习：利用隐式反馈进行推荐。载于数据库系统高级应用国际会议。施普林格，232–247。[6] Hongyu Zhou、Xin Zhou、Zhiwei Zeng、Lingzi Zhang 和 Zhiqi Shen。2023。多模态推荐系统综合综述：分类体系、评估与未来方向。*arXiv*预印本 *arXiv:2302.04473*（2023）。[7] Hongyu Zhou、Xin Zhou、Lingzi Zhang 和 Zhiqi Shen。2023。利用同质图增强二元关系，实现多模态推荐。载于*ECAI 2023*。IOS出版社，3123–3130。[8] Xin Zhou。2022。两个图的故事：冻结和去噪多模态推荐中的图结构。*arXiv*预印本 *arXiv:2211.06924*（2022）。[9] Xin Zhou、Donghui Lin、Yong Liu 和 Chunyan Miao。2023。用于推荐的层细化图卷积网络。载于*2023*年第*39*届*IEEE*数据工程国际会议（*ICDE*）。IEEE，1247–1259。[10] Xin Zhou、Aixin Sun、Yong Liu、Jie Zhang 和 Chunyan Miao。2023。SelfCF：一种简洁的自监督协同过滤框架。*ACM*推荐系统汇刊1，2（2023），1–25。[11] Xin Zhou、Hongyu Zhou、Yong Liu、Zhiwei Zeng、Chunyan Miao、Pengwei Wang、Yuan You 和 FeijunJiang。2023。用于多模态推荐的自举潜在表示。载于*2023*年*ACM*万维网大会论文集。845–854。