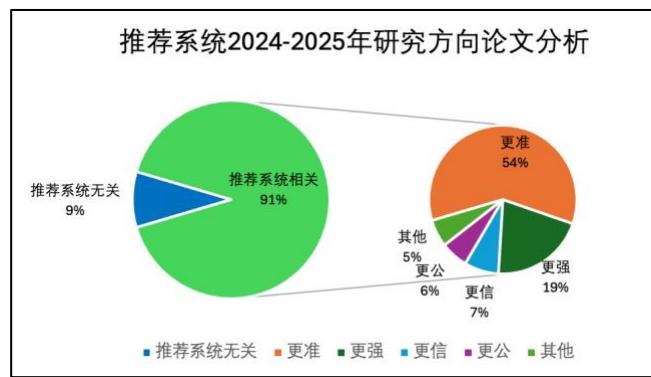


# 西南民族大学本科毕业论文（设计）开题报告

题 目	基于生成式大模型的推荐系统多样性优化与实证研究				
研究类型	<input type="checkbox"/> 理论研究 <input type="checkbox"/> 实验研究 <input checked="" type="checkbox"/> 应用研究 <input type="checkbox"/> 工程设计 <input type="checkbox"/> 软件设计 <input type="checkbox"/> 其他研究				
题目来源	<input type="checkbox"/> 教师科研课题 <input type="checkbox"/> 教师自立题目 <input checked="" type="checkbox"/> 学生自立题目			类 型	<input type="checkbox"/> 设计 <input checked="" type="checkbox"/> 论文
论文（设计）是否在实验、实习、工程实践和社会调查等实践环节中完成				<input checked="" type="checkbox"/> 是 <input type="checkbox"/> 否	
学 院	计算机与人工智能学院	姓 名	刘爱辉	学 号	20223400402 6
年 级	2022	专 业	计算机科学与技术	指导教师	周伟

	<p><b>(一)选题意义</b></p> <p>当今互联网时代下，各类社交媒体和购物等平台都有大量日活用户，推荐系统得以逐渐发展起来。然而，当前（2026年初）推荐系统往往过度追求高准确率，即在海量数据中尽可能匹配与用户兴趣高度相关的内容，这种做法可能已不再适宜。</p> <p>早期推荐算法提升精准度的目标在于从海量信息中快速筛选并推荐相关内容，但目前该方面已趋于完善，以致用户容易陷入“过度精准化”推荐的困境，这反而影响了用户体验。研究者本人亦深切体会到该问题的普遍性与严重性。</p> <p>因此，在提升兴趣匹配度的同时，更应关注用户的身心健康与信息获取的广度。平台推荐机制应当转向“精准推荐与多样性并重”的方向，仅对用户主动搜索的内容保持较高精准度，而在常规推荐中融入多样性策略，以帮助用户减缓“信息茧房”。</p> <p>在大模型技术迅速发展的背景下，如何利用好其对数据深度、潜在信息的映射变得尤为重要。借其突出特性，我将采用生成式推荐策略，来发掘“兴趣推荐+多样化”的潜在可能。</p> <p><b>(二)现状简介</b></p> <p>本研究基于上述现象，系统梳理了2024年初至2025年末推荐系统相关文献，并结合抖音平台（拟作为实证数据来源）的真实推荐数据，开展深入分析。具体工作包括：</p> <ol style="list-style-type: none"><li>首先从推荐系统顶刊数据库中检索了近两年上万篇（10970篇）推荐系统数据，进行清洗、分类、统计后得到如下结果：</li></ol>
--	---



其中与推荐系统相关数据有 9973 篇，无关的有 997 篇（其中包含无关论文和空数据）；与推荐相关的论文大类可划分为：

更准：提升核心效果与体验，提升准确度、推荐性能优化

更强：构建高效鲁棒的架构，推荐架构设计、推荐性能优化(部分板块算法优化)

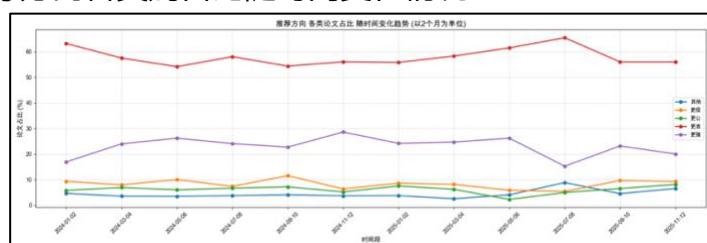
更信：实现透明与可控交互，推荐可解释性

更公：确保公平与社会责任，去偏、公平性、**多样化**

其他：不符合以上 4 类

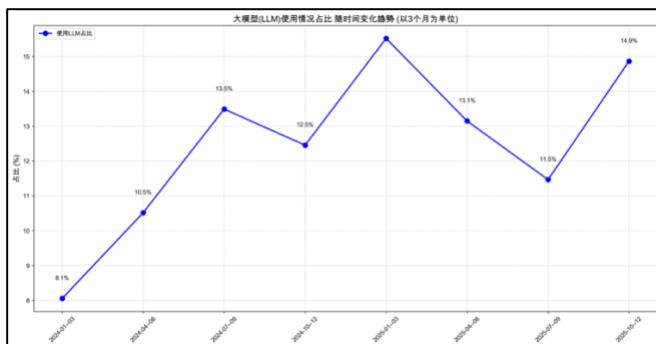
对应类别占比同上图信息（“更准”与“更强”有大多交叉，更倾向“更准”居多），可以看到，“更多” + “更强” 占了近八成，“更公”仅占 6%，而在“更公”里，研究和探讨“多样性”话题的，仅有 52 篇，占比 0.5%。

## 2. 推荐方向各类别占比随时间变化情况：



从时间趋势上看，2025 年较 2024 年研究推荐方向的有所上涨，但探讨和研究“多样化推荐”的，依旧微不足道，希望在 26 年得到工业界与学术界的重视。

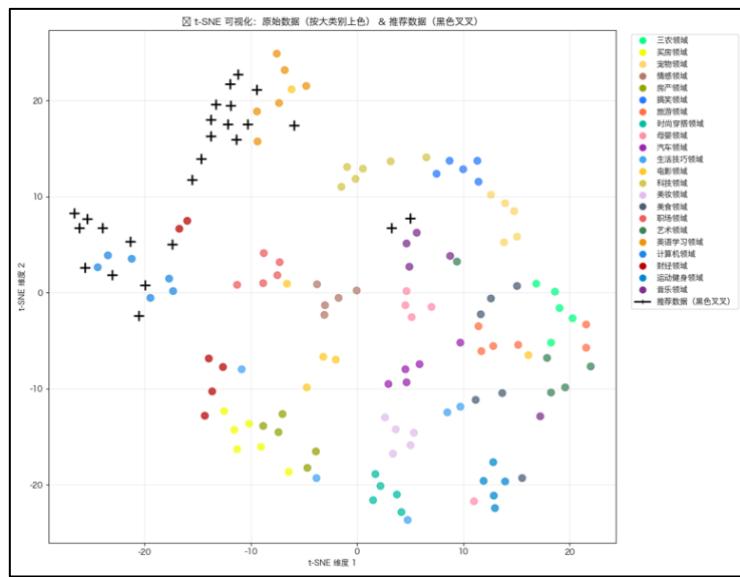
3. 关于利用大模型方案在赋能推荐系统的，25 年较 24 年整体上上涨，趋势如下：



主要研究大致可分为“大模型辅助推荐系统”和“直接利用大模型做推荐系统”。

选题  
意义  
及现  
状简  
介

#### 4. 短视频平台推荐内容真实分布情况：



圆点表示抖音平台的数据分布（目前只尝试了 22 大类），黑色“+”代表推荐给本人的相关视频分布情况。可见，推荐的视频主要分布在“计算机领域”和“英语学习领域”，出现“推荐塌陷”。

至于多样化推荐的研究细节，后续会具体详细进行分析研究。整体上大致分为“前处理”、“中处理”和“后处理”，此次研究主要聚焦于“中处理”，即算法结构层面。

## 摘要

整体上推荐系统的大致背景，多样化主题研究进度，本人采用的方案和结果呈现。

### 一、选题背景与意义

针对当前推荐系统过度追求精准导致的“信息茧房”问题，本研究拟利用生成式大模型探索短视频场景下的多样化推荐策略，以提升推荐内容的广度与用户长期体验。(选题意义部分已说明)

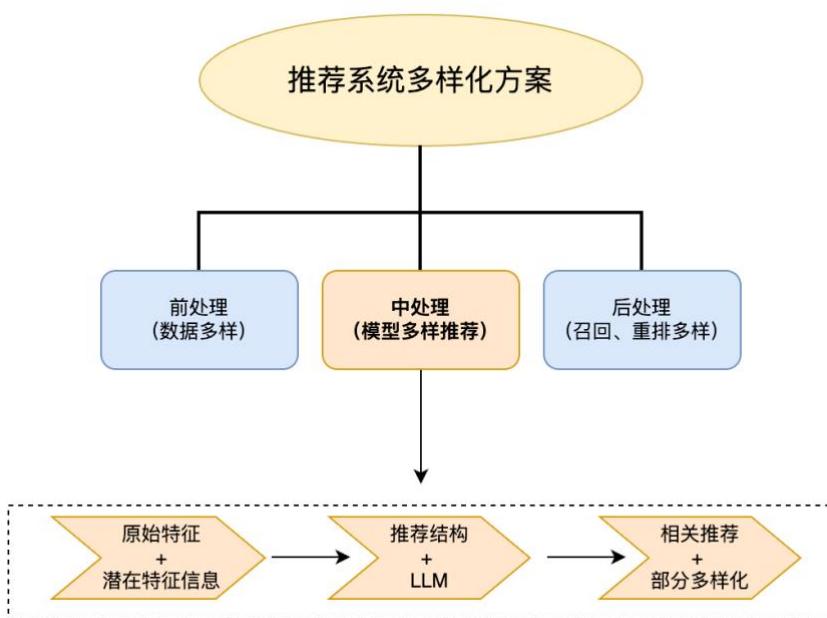
### 二、当前研究现状

目前推荐领域研究仍集中于准确性与性能优化，多样化方向研究占比较低，且利用生成式模型主动创造多样性内容的研究尚不充分。(选题研究现状部分已说明)

### 三、研究目的、内容与结果

旨在设计一种融合生成式的多样化推荐方法，重点研究语义扩展或者其他可行的实验方案，并以真实数据验证效果。

### 四、研究方法与技术路线



本研究遵循“问题量化→方法构建→实验验证”的递进路线。首先，基于公开推荐数据集(如 MovieLens、Amazon Reviews)的用户交互数据，利用嵌入表示与预训练语言模型提取用户与项目的语义特征，并构建“社区内推荐指数”等相关指标以量化同质化强度。随后，设计以生成式大模型为核心的多样化推荐框架：采用大语言模型对用户历史交互序列进行语义扩展与潜在兴趣发散，构建多样性增强的候选集；进而设计轻量级的生成推荐整体化模块，以实现推荐结果的相关性、新颖性与结构多样性。

毕业论文(设计)的主要内容	<p><b>五、预期目标与创新点</b></p> <p>期望构建一个可嵌入实际系统的轻量生成式多样化模块，在保持核心指标稳定的前提下，显著提升推荐结果的多样性。</p> <p><b>六、未来展望</b></p> <p>采用后反馈优化（RL）多样性；</p> <p>数据的构建可能存在一定的局限性，但依旧能表明此现象；</p> <p>缺乏平台真实的线上用户测试；</p> <p>可以考虑的额外方案（实验过程中发现的新现象）；</p> <p>感觉一个重要的生态是能让短视频推荐如此内卷的情况下，脱颖而出的关键。比如微信的社交属性，小红书的经验分享贴，bilibili 的学习长视频，这些不同于其他软件的属性，才是最基础的最牢固的基架。</p> <p><b>七、主要参考文献</b></p> <p>重点参考了近两年推荐系统顶会中关于多样化推荐与大模型应用的论文与报告，以及可能的必要数据集。</p> <p>[1] ZHAO Yuying, WANG Yu, LIU Yunchao, CHENG Xueqi, AGGARWAL Charu C, DERR Tyler. Fairness and Diversity in Recommender Systems: A Survey[J]. ACM Transactions on Intelligent Systems and Technology, 2025</p> <p>[2] Yuying Zhao, Yu Wang, Yunchao Liu, Xueqi Cheng, Charu C. Aggarwal, and Tyler Derr. 2025. Fairness and Diversity in Recommender Systems: A Survey. ACM Trans. Intell. Syst. Technol.</p> <p>[3] Y. Liu and Y. Zhang, "Diversity-Promoting Recommendation With Dual-Objective Optimization and Dual Consideration," in IEEE Transactions on Knowledge and Data Engineering, May 2025</p>
---------------	--

拟解决的问题及思路、方法	<p><b>问题 1：</b>如何发挥出 LLM 丰富的语义特征特征优势，挖掘相关潜在信息（多样化的同时有保留有相关性）？          思路：探索向量空间分布及深层数据分布情况；实验推荐内容生成的“<b>隐层挖掘</b>”能力。</p> <p><b>问题 2：</b>如何在尽可能不降低推荐相关的同时，一定量地推荐一些丰富多样的内容？          思路：依据问题问题 1 的探索研究，发掘潜在多样信息；对长尾数据和热门数据做相对的平衡。</p> <p><b>问题 3：</b>如何系统性地构建算法模块或者算法框架？          思路：探索现有结构和相关研究并做相关适配；根据问题 1 和 2 的相关数据发掘，看能否结合推荐结构，做相关功能的工作。</p> <p><b>问题 4：</b>平台推荐内容同质化信息如何量化（定量）分析？          思路：依据问题 3 的优化目标进行相关的匹配；探索现有多样化、同质化定量分析策略。</p>								
研究进度安排	<table border="1" data-bbox="419 1057 1446 1298"> <tr> <td>2025.12-2026.01</td><td>确定选题及大致目的、分析论文、开题报告</td></tr> <tr> <td>2026.01-2026.02</td><td>实验分析、相关论文研读、整体论文撰写</td></tr> <tr> <td>2026.02-2026.03</td><td>验证相关实验、论文相关细节内容撰写、中期</td></tr> <tr> <td>2026.03-2026.05</td><td>全局论文分析、修改优化、交稿</td></tr> </table>	2025.12-2026.01	确定选题及大致目的、分析论文、开题报告	2026.01-2026.02	实验分析、相关论文研读、整体论文撰写	2026.02-2026.03	验证相关实验、论文相关细节内容撰写、中期	2026.03-2026.05	全局论文分析、修改优化、交稿
2025.12-2026.01	确定选题及大致目的、分析论文、开题报告								
2026.01-2026.02	实验分析、相关论文研读、整体论文撰写								
2026.02-2026.03	验证相关实验、论文相关细节内容撰写、中期								
2026.03-2026.05	全局论文分析、修改优化、交稿								

指导 教师 意见	<p>指导教师签字：</p> <p>年   月   日</p>
答辩 小组 意见	<p>答辩小组组长签字：</p> <p>年   月   日</p>