计算机研究与发展
Journal of Computer Research and Development
DOI: 10.7544/issn1000-1239.202330629
61(7): 1681−1696, 2024

大语言模型驱动的选址推荐系统

高云帆[1] 郁董卿[2] 王思琪[3] 王昊奋[1,3]
[1](上海自主智能无人系统科学中心(同济大学) 上海 201210)
[2](尼斯索菲亚综合理工学院数学学院 法国尼斯)
[3](同济大学设计创意学院 上海 200092)
(gaoyunfan1602@gmail.com)

Large Language Model Powered Site Selection Recommender System

Gao Yunfan[1], Yu Dongqing[2], Wang Siqi[3], and Wang Haofen[1,3]
[1](*Shanghai Research Institute for Intelligent Autonomous Systems* (*Tongji University*), *Shanghai* 201210)
[2](*School of Mathematics*, *Nice Sophia Polytech*, *Nice*, *France*)
[3](*College of Design and Innovation*, *Tongji University*, *Shanghai* 200092)

Abstract Site selection, as a core link in business decisions and urban infrastructure planning, plays a significant role in whether physical stores and urban infrastructure can perform as expected. The existing site selection recommender system's data service orchestration is rather fixed and unable to make timely adjustments to different user needs. Its application scenarios are limited, and the system's flexibility and scalability in human-computer interaction are poor. Recently, large language models (LLMs) such as GPT-4 have demonstrated powerful capabilities in intent understanding, task orchestration, code generation, and tool usage. They can accomplish tasks that traditional recommendation models struggle to balance, providing new opportunities for reshaping the recommendation process and implementing integrated recommendation services. However, site selection recommendation not only shares the common challenges of traditional recommendations but also presents unique challenges due to its reliance on spatial data. Against this backdrop, we propose a LLM-powered site selection recommendation system. Firstly, we expand the scenarios of site selection recommendation and propose a scene recommendation task of finding suitable store types based on location, combining collaborative filtering algorithms and spatial pre-training models. Secondly, we construct a site selection decision engine driven by a large language model. The language model itself has many defects in dealing with space-related tasks, such as the lack of spatial awareness, inability to understand specific locations, and the tendency to fabricate place names and addresses. In this paper, we propose a mechanism for handling spatial tasks within the language model framework. By utilizing geocoding, reverse geocoding, and address resolution tools, we enhance the model's spatial awareness and prevent address fabrication issues. In combination with site selection recommendation models, scenario recommendation models, external knowledge bases, and map visualization, we accomplish diverse tasks in site selection recommendations. We achieve intelligent planning, execution, and attribution for site selection tasks, enhance the interaction experience of spatial service systems, and provide new design and implementation ideas for future AI-driven site selection recommender systems.

收稿日期: 2023-07-31; **修回日期**: 2024-01-04
基金项目: 国家自然科学基金项目(62176185); 信息系统工程全国重点实验室开放基金(PU52221147); 土木工程I类高峰学科建设项目(2022-3-YB-02)
This work was supported by the National Natural Science Foundation of China (62176185), the National Key Laboratory of Information Systems Engineering (PU52221147) and the Top Discipline Plan of Shanghai Universities-Class I (2022-3-YB-02).
通信作者: 王昊奋(carter.whfcarter@gmail.com)

Key words site selection recommendation; recommender system; large language model; agent; geospatial analysis

摘 要 选址作为商业决策和城市基础设施规划的核心环节，对实体店铺、城市基础设施能否发挥预期效用具有重要作用.现有的选址推荐系统数据服务编排较为固定，无法对不同用户需求系统做出及时调整，应用场景受限，人机交互的系统灵活性和可扩展性差.最近，以GPT-4为代表的大语言模型（large language model, LLM）展现出了强大的意图理解、任务编排、代码生成和工具使用能力，能够完成传统推荐模型难以兼顾的任务，为重塑推荐流程、实现一体化的推荐服务提供了新的机遇.然而，一方面选址推荐兼具传统推荐共有的挑战；另一方面，由于其基于空间数据，具有独特的挑战.在这一背景下，提出了大语言模型驱动的选址推荐系统.首先，拓展了选址推荐的场景，提出了根据位置寻找合适店铺类型的场景推荐任务，结合了协同过滤算法和空间预训练模型.其次，构建了由大语言模型驱动的选址决策引擎.语言模型本身在处理空间相关的任务上存在诸多缺陷，例如缺少空间感知能力、无法理解具体位置、会虚构地名地址等.提出了一种在语言模型框架处理空间任务的机制，通过地理编码、逆编码、地名地址解析等工具提升模型的空间感知能力并避免地址虚构问题，结合选址推荐模型、场景推荐模型、外部知识库、地图可视化完成选址推荐中的多样化任务.实现选址任务的智能规划、执行与归因，提升了空间服务系统的交互体验，为未来人工智能驱动的选址推荐系统提供新的设计和实现思路.

关键词 选址推荐；推荐系统；大语言模型；智能代理；地理空间分析

中图法分类号 TP391

选址对于商业、工业和公共设施的成功至关重要.正确的选址可以大大提高运营效率，带来更多的客流，降低运营成本，从而提高业务的成功率[1].零售店铺需要在其目标客户频繁出入的地方设立，充电站需要设立在驾驶者经常经过或停留的地方，以便他们可以方便地充电，而工厂需要靠近其关键供应商和/或目标市场以降低运输成本.在选址过程中，需要综合考虑各种因素，包括但不限于人流量、交通情况、地价、社区发展规划以及周围环境的竞争和协同效应.过于靠近直接竞争者可能会引发潜在客户的分流，而如果过于远离，则可能会错过一些交叉销售的机会.这种复杂的选址决策通常需要深入的市场研究和全面的评估.

传统的选址方法主要依赖于经验规则和人工考察，例如，根据店铺周边的配套设施、人口统计数据、客流分布等进行经验性的总结和判断.这种方法存在一些明显的局限性，例如容易受主观因素的影响，难以在大范围内应用，同时也耗费时间和成本高昂.随着机器学习方法的应用，基于数据驱动的选址方法已开始在连锁企业的选址决策中发挥作用.利用日益丰富的空间数据和算法，克服了传统方法的局限性，提供了一种更高效、准确的方式来进行选址决策.

尽管当前选址推荐系统在效率上已有显著提升，但面对复杂的选址需求，其表现仍然存在不足.首先，现有的选址推荐服务主要针对大型连锁企业[2]，这些企业通常拥有大量的数据和广泛的需求.然而，对于小型企业或单一设施的选址决策来说，常常面临空间数据和历史经验不足的问题.其次，当前选址系统的应用场景相对单一，它们主要针对连锁品牌的新店铺开设，而对其他相关选址服务的覆盖不足.例如，缺乏对推荐点位的归因分析——解释为什么某一位置优于其他位置.此外，不同于从行业品牌出发寻找点位，有时需要根据特定点位推荐合适的行业[3].这种推荐场景对于招商招租和广告牌招租等业务极为重要，但尚未被现有的选址推荐系统重视.

在推荐服务的角度上，当前的选址推荐系统在数据服务编排上相对固化，不能对不同用户需求进行及时的调整，从而限制了应用场景的广度.用户只能依赖于系统预设的服务来满足自己的选址需求.此外，现有系统在人机交互的灵活性和可扩展性方面表现不足.目前，选址推荐系统大多使用可视化大屏作为用户交互的界面，尽管可以向用户展示推荐的点位，但在提供决策辅助服务方面能力有限.在交互形式上还有许多需要改进的地方，这不仅包括提高对不同用户需求的响应能力，而且包括增强人机交互的灵活性和可扩展性，以便提供更全面、更深入的决策辅助服务.

最近，以GPT-4[4]为代表的大语言模型（large language model, LLM）展现出了强大的综合能力，能够生成自然、连贯且相关的文本，使得人机对话更加流畅自然.LLM可以理解人类意图，能够胜任更困难的任务.在复杂的决策环境中，LLM不仅可以借助思

维链，通过逐步推理生成输出结果[5]，还可以调用外部工具[6]，LLM 本身作为系统的推理引擎，赋予其各类外部工具以及自身的长期记忆让其自行生成灵活决策步骤的研究范式也被称为智能代理（agent）.

LLM 展现出了彻底改变推荐系统的巨大潜力[7]. 其在推荐系统中扮演的角色开始逐渐突破传统定位，从简单的编码器、打分器向外扩展延伸，在特征工程乃至推荐流程控制中都发挥重要作用，能够完成用户意图推断、知识延展外推等传统模型难以兼顾的任务，这向重塑推荐流程、实现推荐系统一体化迈进了一步[8]. 最近的研究试图利用 LLM 的力量来增强推荐系统. 文献 [9] 融合传统推荐系统的商品召回能力，以及 LLM 理解和交互能力，发现可以更好地实现交互式推荐和跨域推荐. 文献 [10] 用自然语言描述用户的偏好、意图、任务形式和上下文，并通过预置的指令模板让 LLM 完成推荐任务.

选址推荐系统，不仅面临着传统推荐系统所共有的挑战，而且由于其基于地理空间数据，具有其独特的任务场景特性. 这使得当前 LLM 增强传统推荐系统的方法并不完全适用于选择推荐等空间数据依赖的任务[11]. 由于 LLM 的预训练语料以文本为主，生成方式以预测下一个词元（token）为主，这导致 LLM 很难理解地理空间与地名地址的概念. 对于 LLM 来说，地名地址只是文本，没有具体的地理位置含义. 尽管在学习大量语料后，LLM 可能对经常出现的地址有了更深入的理解，但是对于地址或地理空间有很多名称之外的额外信息如空间距离、位置、属性等，并没有很好地捕获和刻画，导致 LLM 对空间位置的理解是模糊的. 举例来说，LLM 可以明确地指出上海市有几个行政区，但是无法判断它们之间的相对位置关系. 在回答涉及具体地址的问题时，因其“幻觉”也常常会生成虚构的地址. 因此，当我们试图利用 LLM 提升选址推荐系统的服务水平时，会面临更多的挑战.

在考虑到上述背景后，为了使数据驱动的选址推荐系统满足当前日益复杂且不断变化的需求，本文采用了一种基于多源地理数据增强的策略. 这种策略能够在目标源数据稀缺的场景下，为候选品牌的选址扩展提供更多的训练数据. 同时，我们结合了选址分类模型与地理空间预训练模型，在不依赖除兴趣点的（point of interest, POI）以外的空间数据的情况下，提供稳定且可靠的服务，从而大大减少了对标签数据的依赖. 通过基于特征重要性和特征映射的方案，我们能够将选址结果归因到特定的维度上，从而在不同候选地址之间提供直观的对比和理解. 此外，本文还提出了一个面向地址推荐品类的场景推荐任务，这不仅扩展了当前选址推荐的应用场景，而且提高了选址推荐系统在实际应用中的服务能力和效果.

另一方面，本文探索了构建 LLM 与地理空间任务之间连接机制的方法. 通过调用地理编码、逆编码、地名地址解析工具赋予了 LLM 空间感知和位置理解能力. 并作为系统的决策大脑，利用外部知识库为 LLM 提供及时的行业分析和洞见，以帮助用户进一步了解行业发展现状. 对用户的任务进行智能拆解，并根据需求调用选址模型、场景推荐等 API，以满足用户的不同需求. 最后，LLM 对相关信息进行整合和 2 次加工，以提供可解释的选址决策结果，同时也使得决策过程可交互.

本文的主要贡献包 3 个方面：

1）从数据驱动的多角度出发，在数据增强、选址结果可解释、任务扩展等方面对当前选址推荐系统进行了完善，使其更能胜任日益复杂的实际选址推荐场景.

2）提出了一种由 LLM 驱动的选址推荐引擎. 综合利用多种地理编码工具和选址模型实现选址任务的智能规划、执行和归因，提升了 LLM 在地理空间任务上的表现.

3）在广泛的真实数据上验证我们提出的选址推荐和场景推荐方法，并展示了 LLM 驱动的选址推荐的相关案例.

1 相关工作

1.1 选址推荐

近年来，由于城市化发展趋势和城市大数据集的采集便利性[12]，选址推荐作为地理空间人工智能的关键应用场景之一，已引起广泛关注[13]. 文献 [14] 从连锁店铺规模的角度出发，提出了 ChainRec 模型，该模型综合提取地理特征、商业特征和规模特征，以满足不同地区的服务要求和利润优化. 除了商业需求，选址任务也被拓展到了新电源汽车充电站[15]、仓储设施选址[16] 等领域.

然而，在许多场景下，例如连锁企业试图进入新市场或新城市时，由于缺乏历史位置参考，会面临冷启动情况和数据缺失的问题. 针对这一问题，文献 [17] 提出了 CityTransfer 框架，该框架基于协同过滤设计了城市内外双重语义提取和知识迁移方法.

文献 [18] 提出了一种多源多层级选址数据增强方法，通过行业、城市、行政区域和品牌等不同数据源扩充候选品类的选址数据. 文献 [19] 将不同规模区域内的选址问题类比为自然语言翻译任务中的领域不平衡问题，并提出了一种无监督的 Seq2Seq 迁移方法. 文献 [20] 提出了 WANT 框架，设计了一个鉴别器用于减少源城市和目标城市间的分布差异. 该鉴别器与特征提取器一起，通过对抗性学习来学习跨城市的可迁移表征.

除了传统的城市数据和商业数据，多模态数据也被用于进一步提升选址推荐的效果. 文献 [21] 提出了 AR2Net 框架，利用卫星图像、夜间灯光数据以及商业数据之间的交互关系，结合回归和排名损失函数进行训练和评估，以提取与商业地点受欢迎程度相关的特征.

目前大多的选址推荐任务都是面向连锁企业寻找合适的点位，但是却忽略了从点位出发寻找合适品类的需求. 在实际业务中，主要包括商场的招商引资和广告招租等场景. 在早期，文献 [3] 基于社交媒体的签到数据分别提取位置特征和商业特征用于店铺类型推荐. 但是由于签到数据的信息密度较低，实际应用范围较窄.

1.2 智能代理

如 ChatGPT[22] 和 GPT-4[4] 等 LLM 都展现出了强大的复杂推理、意图理解、代码生成等能力[23]. 其中很多能力都被认为在是模型参数达到一定规模才会出现的涌现能力[24]. 越来越多的研究在探索如何利用 LLM 的能力来优化当前的信息系统，这加速了自主智能代理的研究与开发，出现了以 AutoGPT, BabyAGI, AgentGPT, HuggingGPT 为代表的一系列自主代理系统[25].

在以 LLM 驱动的自主代理系统中，LLM 作为代理的大脑，辅以多个关键组件. 研究人员提出智能代理等于大语言模型、记忆、规划技能和工具使用的加和[26]. 文献 [5] 提出了任务矩阵的概念，将基础模型和现有的应用程序结构通过 LLM 组合起来完成数字和物理领域的任务. 文献 [27] 提出了 ChemCrow，它整合了 17 种由专家设计的工具，增强了 LLM 在化学领域的能力. 文献 [28] 提出了 Autonomous GIS 的概念并开发了 LLM-Geo 原型系统，其以 GPT-4 为核心，在没有人工干预的情况下生成 Python 代码执行地理空间数据分析和可视化的任务. 与本文方法相比，LLM-Geo 侧重通过代码生成解决地理数据处理的问题，并没有很好地解决 LLM 在地理空间文本上的幻觉问题. 依赖代码生成也使得 LLM-Geo 在应用范围上存在局限性. 而本文则是聚焦综合调用多种工具，提升 LLM 在空间任务上的感知能力，构建选址决策引擎来实现面向用户的智能选址推荐系统.

2 本文系统框架

本文系统的框架如图 1 所示，共包括 4 个部分：数据层、模型层、应用层和交互层. 其中白色部分为传统选址推荐系统架构，灰色部分本文提出的增强部分.

1)数据层. 多源时空数据作为整个系统的支撑，主要包括兴趣点、轨迹数据、企业数据和用户画像等. 由于多源时空数据异质性的特点，本文将其对应到统一的空间单元中，使用 7 级 Geohash 网格作为最小空间单元，每个网格为 150m×150m(赤道附近). 并通过特征工程，构建高维特征衡量空间特性，例如竞争性、多样性、交通便利性以及附近居民的消费水平.

2)模型层. 单一模型通常难以应对复杂的选址任务. 本文提出的选址推荐系统在技术选型上根据

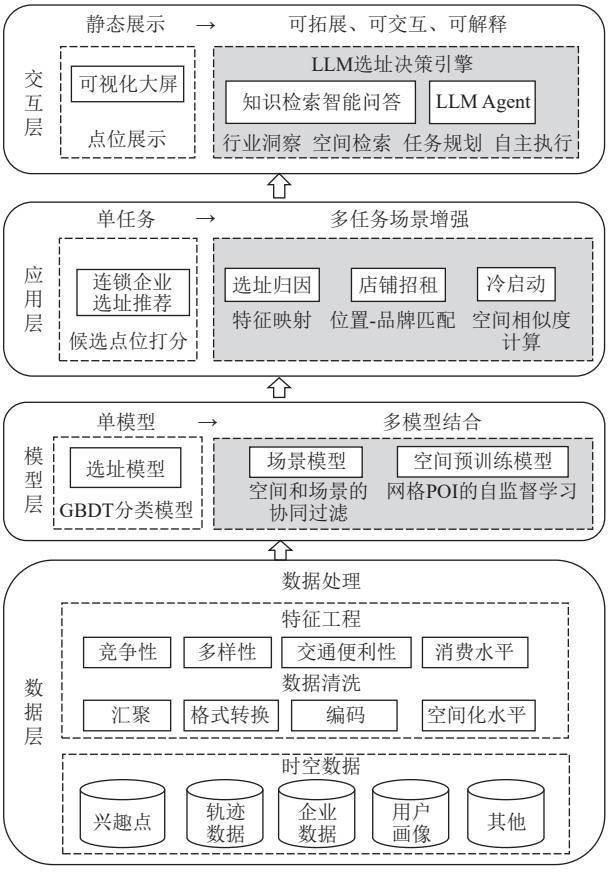


Fig. 1 Illustration of our system framework diagram

图 1 本文系统框架图

不同的业务场景需求，融合了决策树模型，基于协同过滤改进的推荐方法和空间预训练模型，以充分发挥各个模型的特点.

3）应用层. 相较于传统选址推荐系统，我们实现了从单一场景到多场景的拓展. 现有模型通过对地点打分和排序来选址，但得分无法直观揭示地点间的差异. 因此，我们提出基于特征映射的可解释选址归因法，提出面店场景推荐品牌的场景推荐任务. 同时，选址推荐和场景推荐都依赖空间特征，而某些特征（如用户画像和客流量）获取成本高. 因此，我们提出在无法获取相关空间特征时，使用不依赖特征数据的空间预训练模型，并使用基于空间相似度匹配的方法缓解数据缺失导致的冷启动问题.

4）交互层. 在传统的选址推荐系统中，最终的呈现方式大多是将模型预测合适的点位通过大屏展示给用户，用户和系统的交互方式较为单一. 因此，构建知识检索智能问答来提供更加及时的行业信息和建议. 利用 LLM 的工具使用能力，构建选址决策引擎，使得用户可以通过自然语言的交互形式实现自定义的选址推荐任务.

2.1 选址推荐方法

本文提出了一种以历史店铺位置作为正样本来指导选址的方法，方法流程如图 2 所示. 对于具体的品牌或行业，当前店铺位置所在的网格作为正样本，对于每个正样本，在半径为 r 的区域内，随机选择一定数量的非当前行业的 POI 所在的网格作为负样本，其中 r 是一个可以调节的超参数. 将收集到的正负样本点位对应到所处的 7 级 Geohash 网格，根据网格编码和行业从空间网格特征库和行业品牌特征库中抽取对应的特征构成训练集. 我们选择梯度提升决策树（gradient boosting decision tree, GBDT）模型作为分类器，模型选型的依据是相比深度学习模型，GBDT 模型在拥有强大性能的同时，具有更好的可解释性，且离线推理的速度更快，更适合线上服务使用.

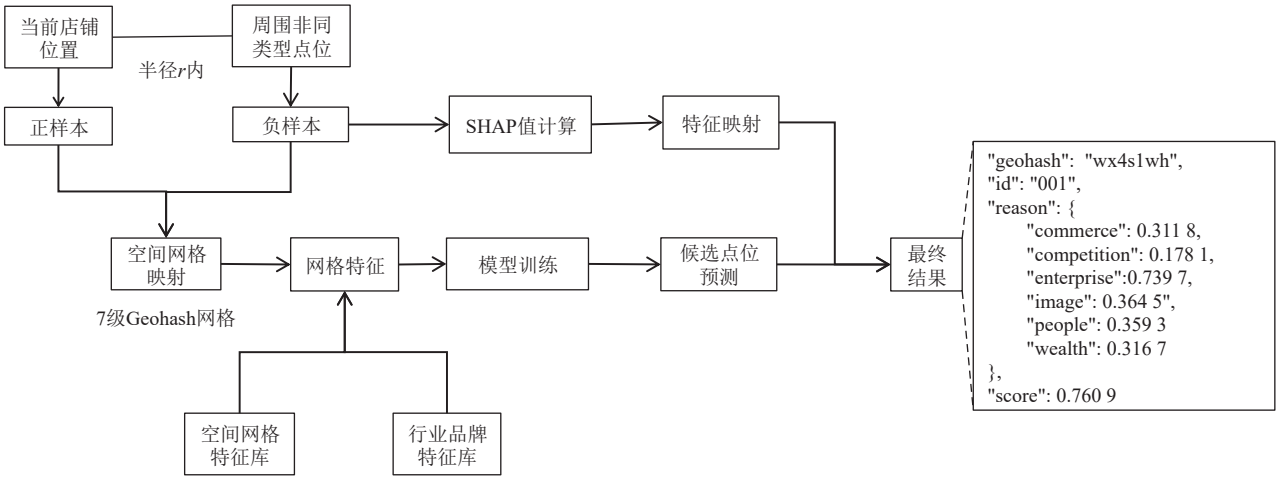


Fig. 2 Site selection recommendation pipeline

图 2 选址推荐流程图

当目标品牌的训练数据不足时，我们通过从 4 种不同的数据源（行业、竞品、行政区域、其他城市）扩充目标源来补充数据. 例如，当品牌进入新市场时，虽然可能缺乏历史数据，但我们可以利用同行业的其他店铺或者来自其他地区的历史数据进行数据增强.

在选址推荐研究中，模型的可解释性一直是一个突出的挑战. 传统的选址推荐模型通常只提供候选推荐位置的评分. 然而，这种分类模型的结果缺乏直观性，对于业务人员来说，很难理解模型的预测结果. 特别是，模型无法直观地解释为何当前位置会被推荐，也无法在多个维度上进行候选位置的比较. 文献 [29] 将店铺位置视为知识图谱中的节点，通过店铺的相对位置和品牌关系，以关系路径的方式对店铺的选址决策进行解释. 然而这种从品牌选址决策角度出发的方法存在 2 个局限：首先可解释路径只能在预先定义的关系选择；其次，只能适合于拥有大量门店的连锁店铺，并与特定品牌强相关，在没有品牌数据的新区域，无法进行直观的解释. 为了解决这些问题，我们使用了 SHAP[30] 模型可解释性方法. 在模型推理时计算每个特征的 SHAP 值，并通过预设的映射表将特征汇聚到具有实际意义的维度. 相比于被应用于传统推荐场景的基于图[31] 或强化学习[32] 的可解释方法，本文方法推理速度快、复杂度低、资源占用少，SHAP 框架可以很好地与 GBDT 模型搭配用于在线服务. 如图 2 所示，在最右

侧的接口实际输出结果中，除了输出模型对候选地址的最终得分，还包括了诸如商业水平、行业竞争水平、企业密集度、人口水平以及周围人群的消费水平等维度的评估，使得对点位的评估更加全面和直观.

2.2 场景推荐方法

当前的选址推荐系统大多没有覆盖从位置出发寻找适合当前点位的任务. 本文提出了一种基于地理位置-品牌的协同过滤方法来进行场景推荐. 如图 3 所示的流程，首先输入地理位置，并从品牌地理位置库中检索近邻位置作为候选地点；然后分别提取输入地点和候选地点对应的人口、环境和地理特征，并利用基于场景的模型计算输入地点和候选地点之间的相似度. 将地理空间和其上的商店类比为经典推荐系统中的用户和物品之间的交互.

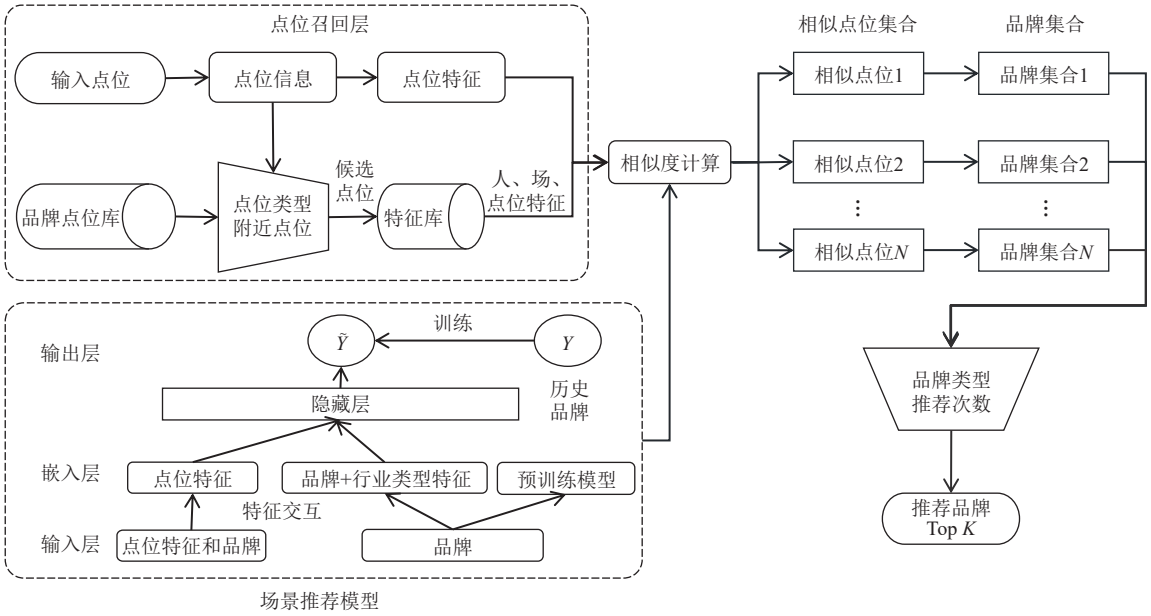


Fig. 3 Scene recommendation pipeline

图 3 场景推荐流程图

为了能够更好地融入品牌特征，我们引入了空间预训练模型 GeoBERT[33]. GeoBERT 将空间网格中的 POI 以最短路径的形式组成序列，将 POI 类型作为 token，通过随机遮盖部分 token 并预测的方式进行空间网格的预训练，实验表明可以在考虑 POI 周围环境的基础上学到 POI 类型的动态表示. 因此，我们在经典的协同过滤算法中加入了由 GeoBERT 生成的 POI 的类别表示. 在此架构中，训练时的标签 Y 表示当前品牌是否在该地点开设过店铺. 在获取到相似地点后，筛选出对应的品牌集合，并根据各品牌的频率和排名权重计算最终的 Top K 推荐. 然而，在实际应用中，涵盖空间网格特征的数据，如用户画像、客流、企业信息、房价、消费水平等，覆盖范围有限. 为了扩大场景推荐的适用范围，在缺少相关特征仅有 POI 数据的情况下时，可以直接使用 GeoBERT 计算空间网格的相似度.

3 选址决策引擎

LLM 展现出强大的语言理解和生成能力，如果将其与其他工具组合使用，则可以实现更复杂的任务. ReAct[34] 和 MRKL(modular reasoning knowledge and language)[35] 系统的成功就证明了这一点，它们将 LLM 与思维链的使用相结合，利用 LLM 的推理能力来控制工具链的执行.

在此背景下，我们提出了一个基于 LLM 的选址决策引擎，系统架构如图 4 所示. 用户可以用自然语言向系统提出选址要求，系统则利用 LLM 的语言理解力解析请求，并调用多个模块如空间数据库、地图引擎、选址模型等来完成选址任务. LLM 既负责解析请求，又可以协调多个模块，最后用自然语言和可视化的形式向用户呈现结果. 相比传统的选址可视化系统，这样的语言交互系统更符合人的思维方式. 同时使得系统任务不再局限于提前预制的服务中，用户可以自主地完成更多定制化的任务.

当用户将指令传入选址决策引擎后，则会经过查询前处理和查询后处理 2 个过程. 在查询前处理中，系统主要执行用户意图理解，根据用户指令从外部知识库中检索相对应的知识. 其次对用户的指令进行任务规划，包括预期使用的工具和对应的参数. 在

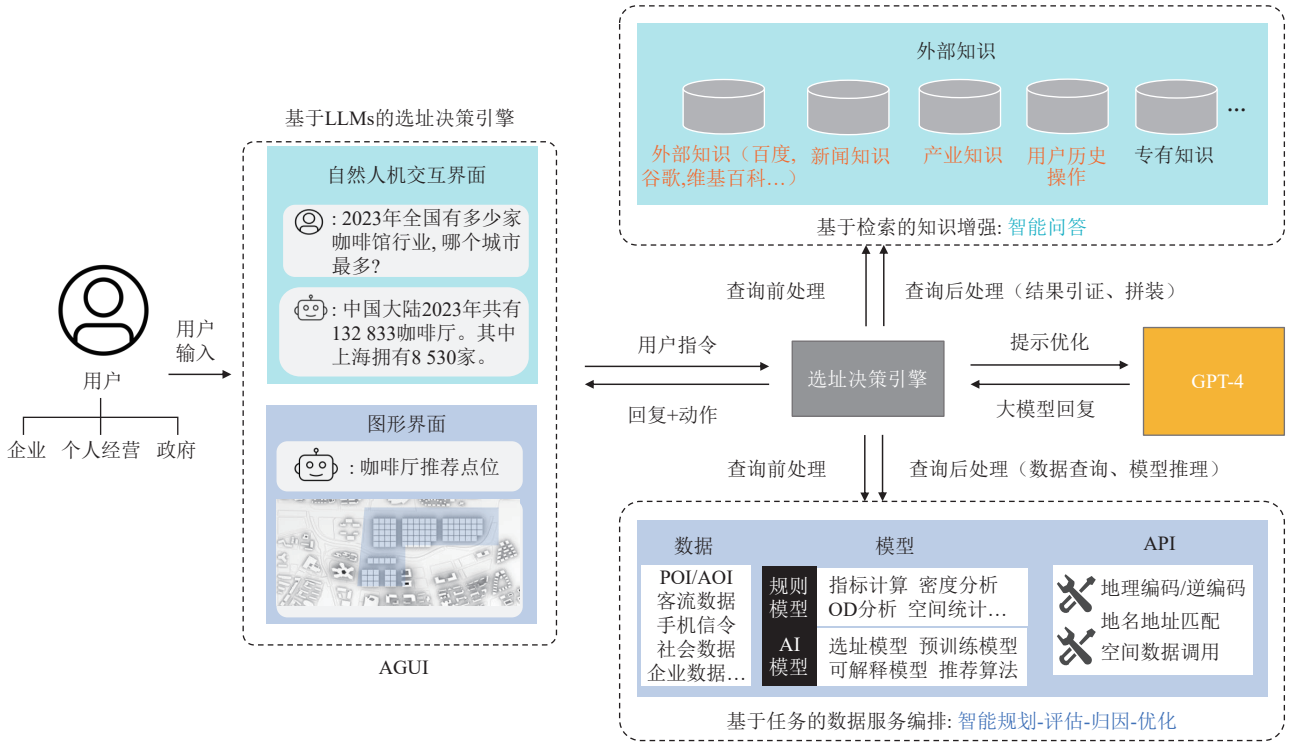


Fig. 4 Site selection recommendation framework

图 4 选址推荐决策框架

查询后处理中，则会根据任务规划，从数据库中查询数据，调用模型. 最后将相关的文本资料、模型结果等进行总结和 2 次处理最终返回给用户. 如果在其中涉及到了具体的点位推荐等行为，则会将相关位置通过可视化的地图显示出来. 在文本中，我们使用了 OpenAI GPT-4 和 MRKL 系统的思路来构建决策引擎，详情请见 3.2 节.

3.1 知识检索智能问答

语言模型受制于训练语料的限制，无法了解最新内容并且缺少专业领域的知识. 通过知识检索的方式可以有效扩充专业知识，提高生成回复的可靠性，有助于减轻“幻觉”问题，并且无需对整个模型进行重新训练，也被称为检索增强生成（retrieval augmented generation, RAG）[36]. 其核心功能在于：1）将用户问题和本地知识进行向量化编码，通过向量相似度实现召回；2）通过 LLM 对用户问题进行意图识别，并对原始答案加工整合.

具体流程包括 3 步：

1）知识库构建. 首先，我们对相关文档进行清洗，将其整理为纯文本格式. 其次，将文档进行分片，划分成若干个独立且较短的段落. 每个段落被视为问答系统的最小记录单元，用于与用户提出的问题进行匹配. 然后，调用语言编码模型，对分片后的段落进行编码. 获取到相应的向量（embedding）后，我们将原始的分片和向量一同以键值对的形式存储到向量数据库中. 为了便于管理和防止在不同行业间出现错误的检索，我们为不同行业设置了不同的命名空间（namespace）.

2）检索. 检索即从知识库获取相关的上下文信息. 如图 5 所示，首先对用户的问题进行信息抽取，通过 slot 填充的方式，确定用户咨询的行业，避免在后续过程中出现行业匹配错误的情况. 其次对用户的问题通过同样的语言模型编码，得到对应的向量表示. 最后获取到与问题相关的 K 个文档分片.

3）对话. 将检索到的相关文档、任务提示以及历史对话组成的增强提示信息交由语言模型进行处理. 这个处理过程包括去重、过滤、摘要、概述、格式处理以及结合历史对话的分析. 每次交互式对话完成后，都会让语言模型对对话历史进行总结，然后将总结结果存储. 当本地搜索的结果不满足要求，例如，相似度过低时，则会基于对话历史让 LLM 自行回答.

3.2 MRKL 和工具调用

尽管 LLM 在多个任务中表现优秀，但仍存在一些固有限制，如无法访问所有相关知识、无法处理特定计算或获取实时信息. 同时，LLM 也难以处理复杂数据或空间数据任务，这是本文想要解决的问题.

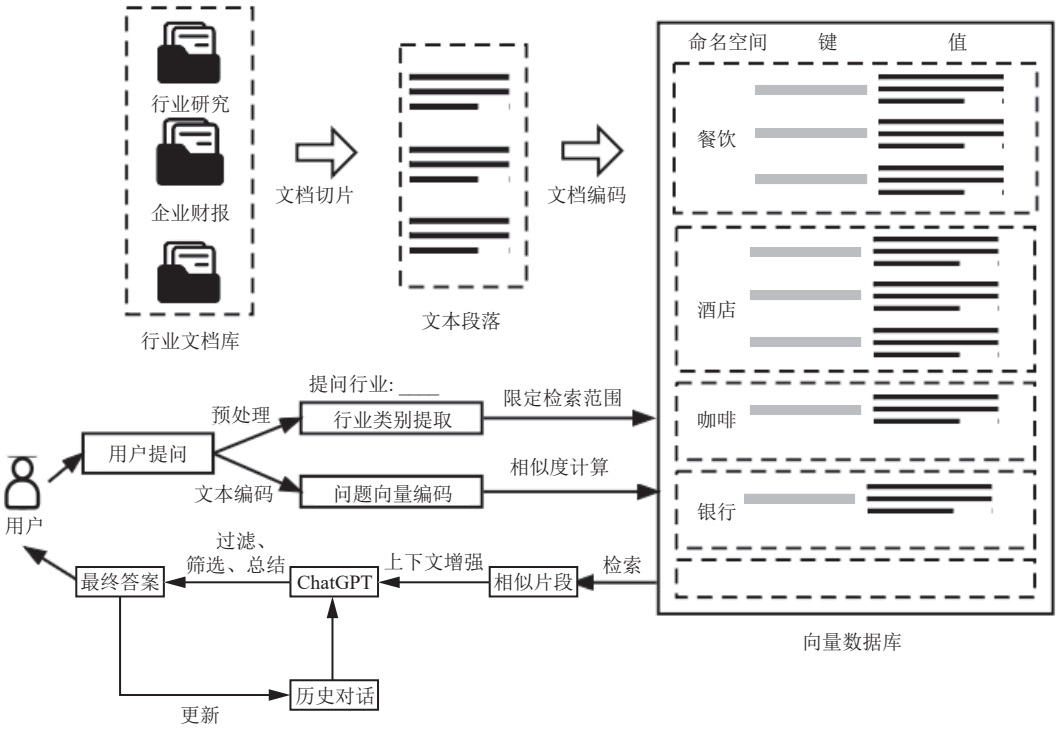


Fig. 5 Retrieval-augmented document QA

图 5 检索增强的文档问答

MRKL 系统是一种模块化的神经符号架构，整合 LLM、外部知识源和离散推理，旨在克服传统语言模型的局限性. 这种架构让神经网络模型处理自然语言输入，外部知识库提供专业知识，而符号系统执行推理任务，实现了知识的高效利用，增加了模型的可解释性和扩展性. MRKL 系统构建于一系列可扩展的“专家”模块以及一个路由器模块，其功能是将每一项自然语言输入有效地路由至最适配的模块. 这些“专家”模块可以是神经网络模型，或者是执行特定任务的工具[35].

由于并没有公开具体的 MRKL 实现流程，因此，在本文中，我们根据 MRKL 架构的思路，结合 ReAct 以及其他相关复现方法，设计了一个选址推荐决策引擎，指令模板如图 6 所示.

整个的决策引擎的核心模板主要包括 4 部分：

1)任务指令说明. 我们首先定义了 LLM 的任务，即扮演一个选址推荐的专家，负责根据用户的需求，在特定的区域为品牌推荐适合的开店位置，以及根据特定的点位推荐合适的品牌. 所有点位都是通过 7 级 Geohash 进行编码，然后将编码输入模型中进行处理. 此外，我们也特别强调了在推荐点位和品牌时，需要提供推荐的理由.

2)工具说明. 在任务指令说明之后，介绍 LLM 可以使用的工具，并说明了这些工具的使用场景和输入输出方式. 以选址推荐模型 site_selection_model 为例子，具体工具描述为：“当你需要做选址推荐，寻找合适的地点开设新店时非常有用. 你需要输入需要进行选址的区域，用上海标准的行政区或街道表示. 你需要确定选址的行业，例如银行业、酒店业、咖啡厅等. 它会以列表的形式返回适合选址的 Top K 个网格(grid)的 Geohash 编码，通常为 7 位字符串. 当你获取到 Geohash 编码后，应使用例如 geohash_decode 等其他工具对其进一步处理.”

3)执行模板. 我们通过模板规定了 LLM 的执行流程：Question, Thought, Action, Action Input, Observation, Final Answer. 这包括 LLM 必须回答的问题、任务规划以及应采取的行动. 执行流程可以重复多次，直到 LLM 认为已经找到最优答案. 为了避免在调用工具解释过程中出现中英文混杂的情况，给后面的输出解析和工具调用造成影响，我们规定这一部分仍然使用英文. 只需要在任务说明中明确指出 GPT 最后必须使用中文回答，则不会影响最后的输出.

4)输出解析. 用户需要解析 LLM 的输出以获取有用的信息. 当检测到“Final Answer”时，用户会接收到最终答案. 在其他情况下，用户会通过正则匹配的方式解析 LLM 的输出，识别出需要调用的工具和工具参数，然后执行具体的工具调用. 此阶段中也包括参数归一化、空白处理和断点重试.

```
你是一个选址方面的专家，你的任务包括，根据用户的要求为特定的店铺推荐合适的位置，以及根据特点的场所推荐合适开始的品牌。你要尽可能完成目标。在工具中，对空间网格进行了7级Geohash编码后，每个网格都有一个唯一的7位字符串id。此外，用户只能理解自然语言，你不能把直接把Geohash返回给用户。在推荐位置和品牌时，请根据已有信息对推荐原因做出合理的解释。最后答案请用中文进行回答。

你可以使用如下工具:
{tools}
site_selection_model:
当您需要做选址推荐，寻找合适的地点开设新店时非常有用。
你需要输入需要进行选址区域，用标准的行政区或街道表示。例如，新江湾城街道
你需要确定选址的行业，例如银行业、酒店业、咖啡厅等。
该工具会以列表的形式返回适合选址的Top N 个网格(grid)的Geohash编码，为7位字符串。
当你获取到Geohash编码后，应使用例如geohash_decode等其他工具对其进一步处理。
不能直接输出Geohash编码。
geohash_decode:
…
Question: the input question you must answer
Thought: you should always think about what to do
Action: the action to take, should be one of [{tool_names}]
Action Input: the input to the action
Observation: the result of the action
… (this Thought/Action/Action Input/Observation can repeat N times)
Thought: I now know the final answer
Final Answer: the final answer to the original input question
These were previous tasks you completed:
Begin!
Question: {input}
{agent_scratchpad}
```

Fig. 6　Prompt temple for site selection engine

图 6　选址决策引擎的指令模板

为了让 LLM 可以更好地处理地名地址数据，本文设计了一种可以在语言模型 Agent 中流通的地名地址处理流程，具体如图 7 所示. 首先，我们让 LLM 提取出用户指令中所涉及到的地名地址. 其次将该地址通过地名地址匹配工具，从对应的标准地名地址库对该地址进行匹配. 在很多情况下，用户输入的地理位置信息并不是标准的，为避免地址上的歧义，需要将其转换成标准的地名. 在获取标准地址后，通过地址编码（geocoding）的形式将具体的地址先转换再对应到 7 级 Geohash 空间网格中. 在系统中其他地理空间工具都以 Geohash 编码作为地址的索引. 因为 Geohash 编码的特性更容易地获取到周围相邻网格.

经过上一步，用户的输入的地址已经被转换为 Geohash 编码，在系统中流动. 但是编码对于用户的含义是不清晰的，语言模型也无法发挥优势. 因此我们使用高德地图地理逆编码的 API，将 Geohash 转换成坐标后，逆编码出当前位置点所处的位置、临近的

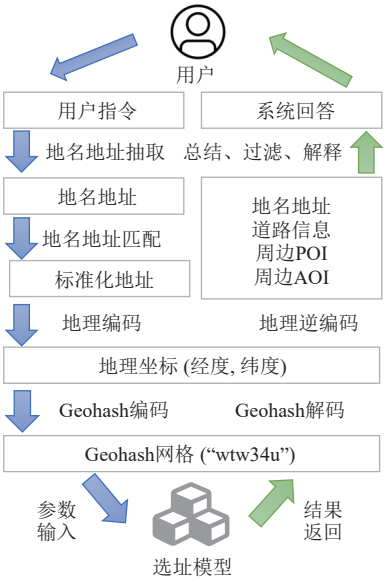


Fig. 7　Address processing pipeline

图 7　地名地址处理流程

道路名称, 以及周围 150m 范围内的兴趣点 POI 和兴趣区域(area of interest, AOI)的描述统计信息, 使得 LLM 能够理解推荐点位的周围环境, 以便整合其他模型输出信息生成最终答案. 为了更好地实现选址功能, 我们定义的工具为:

1)空间数据库. 该工具用户返回地理空间上的点位数据, 例如 POI 和 AOI, 以及空间网格上的特征, 组成其他模型可以使用的数据集. 此外, 通过调用空间数据库, 可以查询到特定区域内所包含的点位, 以及判断点位之间的空间位置关系.

2)空间向量数据库. 利用预训练地理空间模型对空间网格编码并存入向量数据库, 以 Geohash 作为索引. 通过计算向量之间的相似度, 可以快速检索到相应的空间网格.

3)选址推荐模型. 该工具用于制定区域进行选址推荐. 工具的输入为选址的区域、进行选址的行业和需要输出的 Top K 候选点位. 模型的输出位推荐网格的编码和对应维度指标.

4)场景推荐模型. 该工具用于对特定的位置进行品牌推荐. 输入网格编码, 返回 Top K 的推荐品牌类型.

5)地图可视化组件. 该工具用户将其他工具输出的网格位置呈现在可交互地图上. 输入为 Geohash 网格编码.

6)用户. 该工具是人机交互的直接界面, 允许引擎提出问题并期待用户回答, 并期待用户做出回应. 当遇到困难或不确定下一步操作时, LLM 可以请求使用这种工具.

4 实 验

本节将介绍实验数据集、实验相关设置, 然后分别介绍选址推荐和场景推荐的选址结果.

4.1 实验数据

在本文中使用了全国主要城市内多个行业的真实城市数据. 在选址推荐任务中分为行业通用选址和品牌选址. 行业通用将该行业内的所有店铺位置纳入训练集中, 更具有通用性. 具体的行业和对应的数据量(正例数量, 负例数量)分别为: 餐饮(595 592, 1 325 140)、美妆(37 543, 130 536)、咖啡(12 567, 43 304)、银行(168 762, 479 647)、便利店(70 007, 208 924)、酒店(69 188, 203 357)、电影演出(19 139, 73 600)、运动健身(3 017, 9 859)、亲子(38 398, 119 007)、美容塑形(165 865, 399 474)、购物(196 688, 483 924)、学习培训(62 889, 174 678)、园区(20 323, 66 828)、休闲娱乐(125 912, 307 290)、婚庆(3 106, 9 881)、医疗健康(195 864, 465 079). 品牌行业选址以 5 个城市的浦发银行为例, 并根据文献 [18] 中的方法, 使用规模接近的股份制银行对浦发银行的点位进行扩充, 正负样本数据量分别为: 上海(1 422, 2 920)、北京(1 084, 2 682)、广州(616, 1 372)、西安(530, 934)、武汉(662, 1 422). 场景推荐数据集使用上海市范围内数据作为实验数据集, 在过滤掉有效历史点位类型数据后, 共包含 398 526 条正样本和 99 632 条负样本.

4.2 实验设置

选址推荐中方分类模型为 LightGBM[37], 每个网格通过特征提取模块被表示为地理特征向量和用户资料特征向量, 其中范围半径 r 默认为 500 m. 正负样例比例设置为 1∶4, 当负例不足时则全部被使用. 在实验中, 采用贝叶斯超参数搜索, 设定的超参数搜索空间为: *learning-rate*[0.1, 0.2], *max-depth*[6, 8, 10], *min-child-weight*[20, 50, 100], *n-estimators*[30, 50, 100].

在实际推荐应用场景中, 对推荐服务的稳定性和即时性都提出了更高的要求. 在算法和模型方面, 我们从 4 个角度出发:

1)任务相似性. 场景推荐任务是一个较为新颖的任务, 但是其仍与传统的商品推荐具有一定的相似性. 因此, 我们对任务进行抽象, 希望可以借助成熟的传统推荐方法的建模思路. 在我们的任务中, 我们试图找到与特定地点的地理位置特征相匹配的品牌, 这可以看作是用户-商品交互的一个类比, 其中的“点位-品牌”交互类比为“用户-商品”. 在这之前没有特定的技术被应用于场景推荐, 因此我们首先采用较为经典的推荐方法作为对比.

2)数据的适配性. 随着研究的深入和专业化, 推荐方法逐渐引入了如图神经网络[38]、Transformer[39]等复杂结构, 它们也对具体的推荐数据提出了更高的需求, 例如具有图结构或序列结构的数据. 然而, 在场景推荐任务中, 点位和品牌之间并没有天然的图结构或序列结构, 因此, 对于这些复杂算法的适配性较弱.

3)算法的稳定性与性能. 尽管越来越多的推荐方法不断刷新评测数据集的最优结果, 但是实际工业应用中经典推荐方法仍表现出了持续的稳定性和可靠的性能. 这是我们选择它们作为基准的一大原因.

4)线上服务的效率. 场景推荐需要在线上提供及时的服务, 这迫使我们选择一种既简洁又高效的方法. 这也阻碍了我们使用更复杂的算法结构.

基于以上考虑，我们以 NeuralCF[40] 算法的架构为核心，融合地理空间特征和 GeoBERT 预训练模型的知识，将空间网格中的 POI 分布作为上下文信息，拼接上具体的 POI 表征. 基线选用了原始的 NeuralCF，NeuralCF（GMF），Wide&Deep[41].

中间隐藏层层数分别测试 [3, 5, 8, 16] 并选取最优结果. 在所有实验中，训练集和测试集按 8∶2 分割，随机数种子为 42.采用准确率、召回率和 *F*1-score 来评估结果.

与本文使用的 GeoBERT 模型较为相似的是 POIBERT[42]. 二者都使用了 POI 数据和 BERT 的架构用于预训练，各自领域内有独特贡献，但是二者在任务定义、输入数据和训练数据方面都存在差异. 具体来说 POIBERT 的主要任务是面向用户的 POI 序列推荐，预测下一步将会到达的点位. 它依赖于用户的轨迹数据，这些数据来自于图片分享网站 Flickr，通过用户上传照片时的 GPS 坐标获取用户的轨迹数据. 这些数据集包含 7 个海外城市的 5 654 个用户. 相比之下，GeoBERT 的目标是学习通用的地理空间表示，这使得它可以应用于多种下游任务.GeoBERT 使用地理空间网格内的全部 POI 数据，这些数据分布在平面上，并无有序序列规则，是通过不同的方式手动序列化的. 其训练数据是通过高德 API 获取的全国约 30 个城市的 POI 数据，共计约 1 700 万个点位数据. 因此，我们使用了 GeoBERT 用于获取动态的 POI 信息，也并未将 POIBERT 作为实验比较对象.

4.3 实验结果与分析

通用行业的选址结果如图 8 所示，可以看到各个行业的选址分类的准确率都集中在 70%~90% 之间，说明选址模型达到预期效果. 通用选址模型的意义在于，即使没有指定具体的品牌以及对应的品牌数据，依然可以通过行业模型获取到较为稳定的选址效果. 不同的行业之间的选址效果依然存在差异. 例如便利店的准确率为 76.67%，购物为 71.86%，表现相对较弱，而银行（89.63%）、园区（87.49%）则表现较好，可以推测出当行业的位置具有较明显的特性时，其选址效果更好. 而像广泛存在的便利店一样，对选址没有特殊要求，行业选址表现则较弱. 表 1 展示了浦发银行在 5 个城市中的选址推荐结果，可以发现其效果都能超过行业通用选址，说明了当选址范围限定到具体某一品牌并可以获得对应数据时，则可以获得更好的选址效果.

场景推荐的实验室结果如表 2 所示，其中经典的 NeuralCF 和 Wide&Deep 模型作为基线. 其中冷启动场景下是只使用 GeoBERT 中预训练后的 POI 表示，没有依赖额外的地理空间特征. 通过构造地理空间特征，本文方法已经能够取得较好的效果，在拼接上预训练 POI 向量后，性能得到小幅度提升，准确率接近 83% 的准确率. 综上，说明了本文将空间网格上的历史开店行为类比成传统推荐中的用户与产品的交互逻辑是可行的，可以通过传统推荐方法实现基于位置的行业类型推荐.

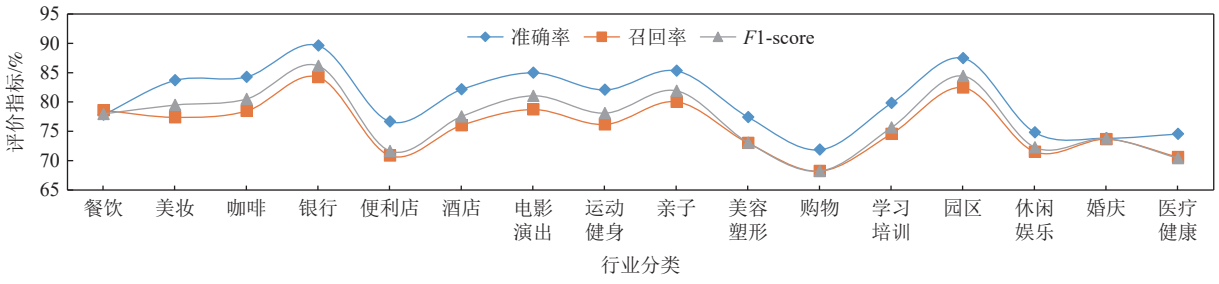


Fig. 8 Industry site selection recommendation results

图 8 行业选址推荐结果

Table 1 SPD Bank Site Recommendation Results

表 1 浦发银行选址推荐结果 %

城市	准确率	召回率	*F*1-score
上海	90.27	91.31	90.74
北京	91.01	91.52	91.26
广州	83.01	85.80	84.18
西安	92.10	93.43	92.68
武汉	84.70	82.86	83.55

Table 2 Scene Recommendation Experimental Results

表 2 场景推荐的实验结果 %

模型	准确率	召回率	*F*1-score
Wide&Deep	80.79	79.66	79.47
NeuralCF（GMF）	75.09	74.95	74.91
本文方法（冷启动）	77.24	65.97	71.16
本文方法（只有特征）	**82.97**	82.73	82.69
本文方法（合并）	82.96	**82.82**	**82.80**

注：黑体数值表示最高的实验结果.

5 案例研究

本节将介绍系统中智能问答、选址推荐和场景推荐的具体使用案例并和原生 GPT-4 的回答进行了对比.

5.1 知识检索增强的智能问答案例

图 9 展示了系统中关于知识检索增强的智能问答的案例. 在实际的拓店选址的任务场景中，无论是连锁企业，还是独立经营店铺，对于是否要在某区域开设一家新的店铺以及开设地点选择的问题，都是一项挑战，不仅受到宏观经济因素的影响，而且与特定行业的发展状况紧密相关. 因此，对于预计开设店铺的商家来说，了解和参考行业发展报告以及行业数据是非常重要的. 基于 LLM，结合给定的业内发展报告与行业数据可以实现快速从大量的行业报告和行业数据中找到与用户查询相关的信息，然后结合 LLM 的生成能力，以智能问答的形式为用户提供详尽而精准的回答. 例如，在案例中，系统根据相关文档阐述了今年美妆、医美和珠宝连锁行业的发展情况，并对各个行业的推荐程度做出了排序. 通过文档快速检索的能力，商家不仅可以快速高效地了解到各个行业的发展情况，还可以了解到行业的风险因素，这对于他们制定开店策略提供了重要的参考. 而原生 GPT-4 则受限于训练语料，无法提供及时的信息.

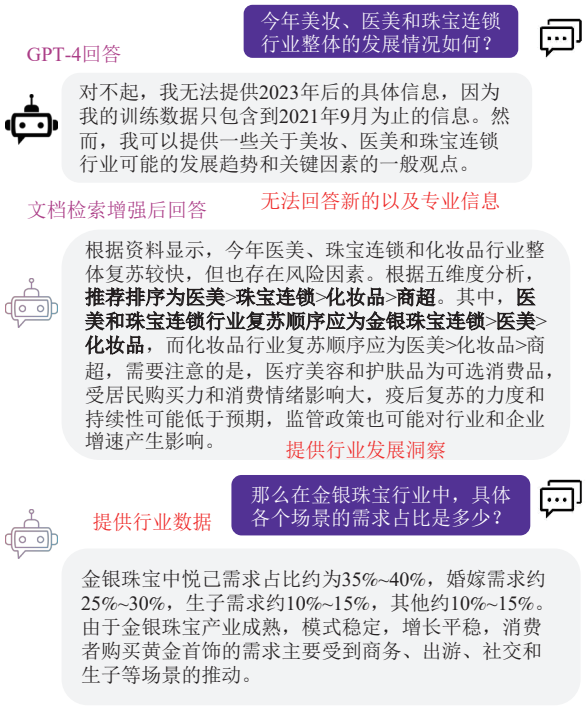


Fig. 9 Illustration of document QA

图 9 文档问答展示

5.2 选址推荐案例

图 10 展示了基于 GPT-4 的选址问答结果与 LLM 驱动系统构建的选址推荐的一组对比. 在直接使用 GPT-4 的问答结果中可以看到多处明显的错误.

首先用户要求在新江湾城街道开设一家咖啡厅，但 GPT-4 推荐的五角场商圈并不属于新江湾城街道而是属于五角场街道，其中出现了事实性错误. 其次 GPT-4 混淆了和街道同名的“新江湾城”小区. 最后由于其本质为语言生成模型，从而输出了“新江湾城商圈”“星光路”“新江湾城中心”等并不存在的虚假地名.GPT-4 的回答充分说明了单一的 LLM 在涉及空间信息任务上的缺陷.

在 GPT-4 下方展示了选址决策引擎在选址任务上的表现. 首先在查询前处理中根据用户给予的信息，从中识别到地址元素，获悉“杨浦区”为“district”

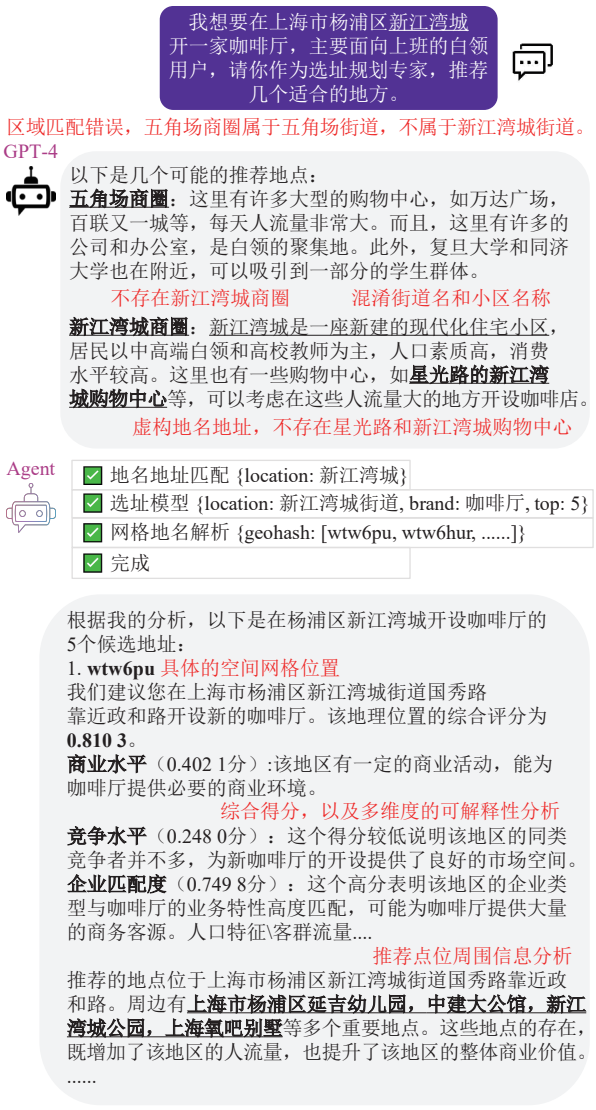


Fig. 10 Illustration of site selection recommendation

图 10 选址推荐展示

实体、“新江湾城”为“neighbourhood”实体，并调用了地名地址匹配工具，根据获悉的实体信息，在标准行政区划地址库中匹配到“新江湾城街道”. 随后在新江湾城街道的网格候选范围内进行选址推荐应用的调用，调用参数涉及用户需要开设店铺的类型（如酒店、银行、便利店等）、区域范围内推荐结果最高的点位的个数（默认为 5）. 在执行完选址模型后，得到模型返回适合开始店铺空间网格的 Geohash 编码、对网格的综合评分以及对应的 6 个维度的可解释性指标. 然而网格的编码对用户是没有实际意义的，用户并不能直观地了解 Geohash 编码所对应的具体位置. 随后 LLM 调用了网格地名解析，查询 Geohash 网格对应的具体文本地名地址和周围环境信息. 最后 LLM 将上述流程获取到的信息进行总结并返回给用户，实现了具体点位的推荐和多维度可解释性分析.

5.3 场景推荐案例

场景推荐的展示案例如图 11 所示. 在原始 GPT-4 的回答中，由于 LLM 并不能理解具体的位置信息，同时也无法获取周边的环境信息，只能给出非常笼统的建议. 因此无法满足用户的场景推荐的实际需求.

在选址决策引擎的流程中，首先通过了地名地址编码工具，将用户输入的地址转换成坐标并映射到空间网格中. 其次，将网格的 Geohash 编码输入到场景推荐模型中，模型会返回推荐的前 3 个行业品类. 最后同样地，通过网格地名解析的工具查询网格对应的具体文本地名地址和周围环境信息.LLM 利用自身的语言能力功底，通过整合上述信息生成对应的回答. 从回答中可以看到，LLM 首先说明了当前地址的周围环境和特点，并结合模型返回的结果推荐了适合当前点位的行业类别.

5.4 讨 论

从 5.1~5.3 节所述的案例中可以发现，对比原生 GPT-4，本文提出的选址决策引擎可以从多个维度提升选址服务水平. 除此之外，本节将本文与其他主流大模型驱动使用外部工具的方法进行对比，以 HuggingGPT 和 AutoGPT 为例.

首先，MRKL 系统并非一种具体的模型方法，而是一种模块化的神经符号架构设计思路. 在广义上，当前许多利用 LLM 调用外部工具的方法都可以归入 MRKL 的范畴.

HuggingGPT 的核心逻辑是 LLM 的自计划和执行.LLM 作为一个控制器，用于管理和组织众多专家模型，如 HuggingFace 中的文生图、图生文、目标检测等模型. 然而，HuggingGPT 的整体流程过于繁琐，

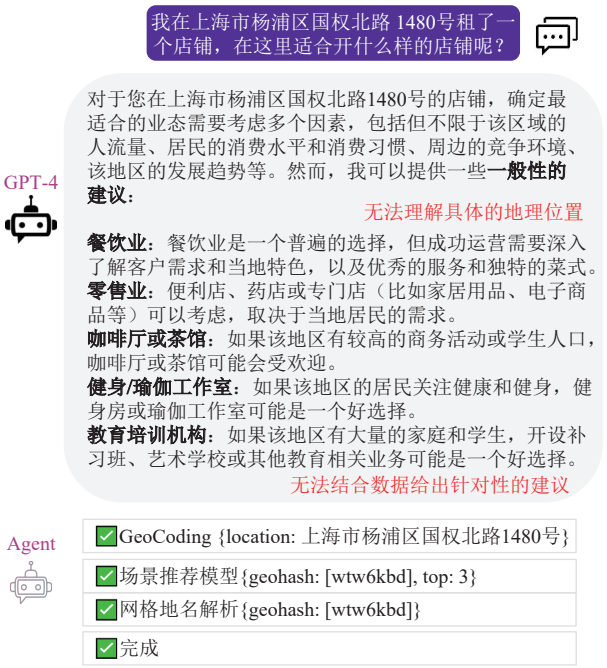


Fig. 11 Illustration of scene recommendation

图 11 场景推荐展示

会消耗大量时间，降低整个系统的服务效率. 而在本文的应用场景中，需要像 HuggingGPT 一样协调众多模型.

AutoGPT 则通过提示设计，给 LLM 增加了更加动态的决策能力，更加灵活和通用. 然而，AutoGPT 的任务链路过长，需要经过十几步甚至几十步的迭代，这对时间和 token 的消耗都很大. 而且，AutoGPT 在很多情况下非常依赖上一步给出的答案，如果和上一步检索出来的答案并不是非常契合，则会导致模型在计划层面出现偏差以及将问题复杂化. 在实际使用 AutoGPT 完成选址推荐的尝试中，AutoGPT 首先会给出较为通用和全面的选址思路，例如做出市场调研，评价当前点位的可达性、可视性和潜在客

流量,提供详细的报告和数据驱动的见解等.然而,在随后进一步检索信息过程中,由于结果较为杂乱,这使其判断,尽管已经获取到了相关场景的部分信息,但是仍需要查找更具体的信息.至此 AutoGPT 陷入了不断检索更多信息的循环中,直至停留在它希望用户可以提供具体网站地址供它检索更多的信息,致使整个任务没有实质性的进展.我们认为 AutoGPT 是对通用人工智能(artificial general intelligence, AGI)的一个重要探索,但是仍处于较为早期的阶段.AutoGPT 的灵活性所带来的相当不确定性使其当前难以在实际场景中应用.而我们的目标是在一个更加可控的范围内提升当前系统的交互服务能力,因此我们当前的方法相比 HuggingGPT 和 AutoGPT 更加合适.

6 总结与展望

本文提出了一个 LLM 驱动的选址推荐系统.对当前推荐系统的应用场景进行了拓展,展示了如何通过外部工具来让 LLM 在地理空间任务中理解具体位置同时避免虚构地名地址.将大语言整合到选址推荐系统中,为用户提供更加智能、灵活和全面的选址解决方案.用户不再被固定的任务场景限制,可以通过交互的形式获取行业信息、模型分析结果和可解释的选址归因.期待能够推动 LLM 等新一代人工智能技术在地理空间任务中的应用与落地.在未来工作中,一方面,我们期待在系统效率方面有所提升,完全依赖 LLM 的调度与执行时间成本较高,充分结合大模型、小模型和引擎等工具,在保证系统自主性的同时提高整体任务的执行效率.另一方面,我们希望能够将更多模态的地理数据引入到系统中,进一步扩展任务覆盖范围.

作者贡献声明: 高云帆负责算法思路、实验方案设计、选址引擎的构建和论文撰写;郁董卿负责完成选址推荐和场景推荐实验;王思琪负责论文的撰写、修改以及可视化系统的搭建;王昊奋负责提出指导意见并修改论文.

参 考 文 献

[1] Liu Yan, Guo Bin, Li Nuo, et al. DeepStore: An interaction-aware wide&deep model for store site recommendation with attentional spatial embeddings[J]. IEEE Internet of Things Journal, 2019, 6(4): 7319−7333

[2] Ge Dongdong, Hu Luhui, Jiang Bo, et al. Intelligent site selection for bricks-and-mortar stores[J]. Modern Supply Chain Research and Applications, 2019, 1(1): 88−102

[3] Yu Zhiwen, Tian Miao, Wang Zhu, et al. Shop-type recommendation leveraging the data from social media and location-based services[J]. ACM Transactions on Knowledge Discovery from Data, 2016, 11(1): 1−21

[4] OpenAI. GPT-4 technical report[J]. arXiv preprint, arXiv: 2303.08774, 2023

[5] Wei J, Wang Xuezhi, Schuurmans D, et al. Chain-of-thought prompting elicits reasoning in large language models[J]. Advances in Neural Information Processing Systems, 2022, 35: 24824−24837

[6] Liang Yaobo, Wu Chenfei, Song Ting, et al. Taskmatrix. AI: Completing tasks by connecting foundation models with millions of apis[J]. arXiv preprint, arXiv: 2303.16434, 2023

[7] Fan Wenqi, Zhao Zihua, Li Jiatong, et al. Recommender systems in the era of large language models [J]. arXiv preprint, arXiv: 2307.02046, 2023

[8] Lin Jianghao, Dai Xinyi, Xi Yunjia, et al. How can recommender systems benefit from large language models: A survey[J]. arXiv preprint, arXiv: 2306.05817, 2023

[9] Gao Yunfan, Sheng Tao, Xiang Youlin, et al. Chat-Rec: Towards interactive and explainable LLM-augmented recommender system[J]. arXiv preprint, arXiv: 2303.14524, 2023

[10] Zhang Junjie, Xie Ruobing, Hou Yupeng, et al. Recommendation as instruction following: A large language model empowered recommendation approach[J]. arXiv preprint, arXiv: 2305.07001, 2023

[11] Mai Gengchen, Huang Weiming, Sun Jin, et al. On the opportunities and challenges of foundation models for geospatial artificial intelligence[J]. arXiv preprint, arXiv: 2304.06798, 2023

[12] Song Xuan, Gao Yunjun, Li Yong, et al. Spatial data intelligence: Concept, technology and challenges[J]. Journal of Computer Research and Development, 2022, 59(2): 255−263 (in Chinese)
(宋轩, 高云君, 李勇, 等. 空间数据智能: 概念、技术与挑战[J]. 计算机研究与发展, 2022, 59(2): 255−263)

[13] Lin J, Oentaryo R, Lim E P, et al. Where is the goldmine? Finding promising business locations through Facebook data analytics[C]// Proc of the 27th ACM Conf on Hypertext and Social Media. New York: ACM, 2016: 93−102

[14] Li Jing, Guo Bin, Wang Zhu, et al. Where to place the next outlet? Harnessing cross-space urban data for multi-scale chain store recommendation[C]//Proc of the 2016 ACM Int Joint Conf on Pervasive and Ubiquitous Computing. New York: ACM, 2016: 149−152

[15] Ghodusinejad M H, Noorollahi Y, Zahedi R. Optimal site selection and sizing of solar EV charge stations[J]. Journal of Energy Storage, 2022, 56: 105904

[16] Saha A, Pamucar D, Gorcun O F, et al. Warehouse site selection for the automotive industry using a fermatean fuzzy-based decision-

making approach[J]. Expert Systems with Applications, 2023, 211: 118497

[17] Guo Bin, Li Jing, Zheng V W, et al. Citytransfer: Transferring inter- and intra-city knowledge for chain store site recommendation based on multi-source urban data[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 1(4): 1−23

[18] Gao Yunfan, Han Dong, Wang Haofen, et al. GeoGTI: Towards a general, transferable and interpretable site recommendation[C]// Proc of the 2022 Int Conf on Web Information Systems and Applications. Berlin: Springer, 2022: 559−571

[19] Yabe T, Tsubouchi K, Shimizu T, et al. Unsupervised translation via hierarchical anchoring: Functional mapping of places across cities[C]//Proc of the 26th ACM SIGKDD Int Conf on Knowledge Discovery & Data Mining. New York: ACM, 2020: 2841−2851

[20] Liu Yan, Guo Bin, Zhang Daqing, et al. Knowledge transfer with weighted adversarial network for cold-start store site recommendation[J]. ACM Transactions on Knowledge Discovery from Data, 2021, 15(3): 1−27

[21] Xu Yanan, Shen Yanyan, Zhu Yanmin, et al. AR2Net: An attentive neural approach for business location selection with satellite data and urban data[J]. ACM Transactions on Knowledge Discovery from Data, 2020, 14(2): 1−28

[22] Ouyang Long, Wu J, Jiang Xu, et al. Training language models to follow instructions with human feedback[J]. Advances in Neural Information Processing Systems, 2022, 35: 27730−27744

[23] Chowdhery A, Narang S, Devlin J, et al. Palm: Scaling language modeling with pathways[J]. arXiv preprint, arXiv: 2204.02311, 2022

[24] Wei J, Tay Y, Bommasani R, et al. Emergent abilities of large language models[J]. arXiv preprint, arXiv: 2206.07682, 2022

[25] Shen Yongliang, Song Kaitao, Tan Xu, et al. HuggingGPT: Solving AI tasks with chatgpt and its friends in huggingface [J]. arXiv preprint, arXiv: 2303.17580, 2023

[26] Weng Lilian. LLM-powered autonomous agents [EB/OL]. (2023-06-23) [2023-06-30] .https://lilianweng.github.io/posts/2023-06-23-agent/

[27] Bran A M, Cox S, White A D, et al. ChemCrow: Augmenting large-language models with chemistry tools [J]. arXiv preprint, arXiv: 2304.05376, 2023

[28] Li Zhenlong, Ning Huan. Autonomous GIS: The next-generation AI-powered GIS [J]. arXiv preprint, arXiv: 2305.06453, 2023

[29] Liu Yu, Ding Jingtao, Li Yong. Knowledge-driven site selection via urban knowledge graph [J]. arXiv preprint, arXiv: 2111.00787, 2021

[30] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[J]. Advances in Neural Information Processing Systems, 2017, 30: 4765−4774

[31] Shimizu R, Matsutani M, Goto M. An explainable recommendation framework based on an improved knowledge graph attention network with massive volumes of side information[J]. Knowledge-Based Systems, 2022, 239: 107970

[32] Guo Lei, Zhang Jinyu, Chen Tong, et al. Reinforcement learning-enhanced shared-account cross-domain sequential recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2022, 35(7): 7397−7411

[33] Gao Yunfan, Xiong Yun, Wang Siqi, et al. GeoBERT: Pre-training geospatial representation learning on point-of-interest [J]. Applied Sciences, 2022, 12(24): 12942

[34] Yao Shunyu, Zhao J, Yu Dian, et al. ReAct: Synergizing reasoning and acting in language models[J]. arXiv preprint, arXiv: 2210.03629, 2022

[35] Karpas E, Abend O, Belinkov Y, et al. MRKL systems: A modular, neuro-symbolic architecture that combines large language models, external knowledge sources and discrete reasoning [J]. arXiv preprint, arXiv: 2205.00445, 2022

[36] Lewis P, Perez E, Piktus A, et al. Retrieval-augmented generation for knowledge-intensive NLP tasks [J]. Advances in Neural Information Processing Systems, 2020, 33: 9459−9474

[37] Ke Guolin, Meng Qi, Finley T, et al. LightGBM: A highly efficient gradient boosting decision tree[J]. Advances in Neural Information Processing Systems, 2017, 30: 3146−3154

[38] Liu Han, Wei Yinwei, Yin Jianhua, et al. HS-GCN: Hamming spatial graph convolutional networks for recommendation[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(6): 5977−5990

[39] Li Jiacheng, Wang Ming, Li Jin, et al. Text is all you need: Learning language representations for sequential recommendation [J]. arXiv preprint, arXiv: 2305.13731, 2023

[40] He Xiangan , Liao Lizi, Zhang Hanwang, et al. Neural collaborative filtering [C] //Proc of the 26th Int Conf on World Wide Web. New York: ACM, 2017: 173−182

[41] Cheng H T, Koc L, Harmsen J, et al. Wide & deep learning for recommender systems [C] //Proc of the 1st Workshop on Deep Learning for Recommender Systems. New York: ACM, 2016: 7−10

[42] Ho N L, Lim K H. POIBERT: A transformer-based model for the tour recommendation problem [C] //Proc of the 2022 IEEE Int Conf on Big Data (Big Data). Piscataway, NJ: IEEE, 2022: 5925−5933



Gao Yunfan, born in 1997. PhD candidate. Student member of CCF. His main research interests include spatial-temporal data ming, spatial-temporal knowledge graph, and large language model.

高云帆, 1997 年生. 博士研究生. CCF 学生会员. 主要研究方向为时空数据挖掘、时空知识图谱、大语言模型.



Yu Dongqing, born in 1997.Master candidate. His main research interests include natural language processing and recommender system.

郁董卿, 1997 年生. 硕士研究生. 主要研究方向为自然语言处理、推荐系统.



Wang Siqi, born in 1993. PhD candidate. Student member of CCF. Her main research interests include urban computing, generative design, and AI-generated content.

王思琪, 1993 年生. 博士研究生. CCF 学生会员. 主要研究方向为城市计算、生成式设计、人工智能生成内容.



Wang Haofen, born in 1982. PhD, professor. Senior member of CCF. His main research interests include knowledge graph, natural lanugae processing, and large language model.

王昊奋, 1982 年生. 博士, 研究员. CCF 高级会员. 主要研究方向为知识图谱、自然语言处理、大语言模型.