

基于多任务学习的商业选址推荐方法的研究

陈麒先

北京航空航天大学，计算机学院

摘要 商业选址是一类重要的投资决策问题。其重要性主要体现在投资的长期性、固定性以及经济效益的决定性上。在传统的商业选址问题中，通常的考量因素往往涵盖了地域、交通、竞争压力以及人流量等方面。在这种情况下，投资者的经验以及数据信息来源的有效性将起到决定性的作用。随着移动互联网时代的到来，越来越多的商业应用，如美团、大众点评等渐渐走入人们的生活。这其中蕴含着巨大潜在的商业价值有待挖掘，尤其对于商业选址这类重要的问题而言，数据所提供的参考信息已然成为大数据时代的选址利器。

为此，本文对深度多任务机器学习模型在商业选址推荐方面的应用进行了考察和研究。通过对多个商业选址的多个决定因素进行分析，并对使用多任务学习在商业领域的应用情况进行了综述。从大众点评的用户评论中采集到的大量用户评论数据表明，通过提取出与商场有关的内部与外部特征，应用关联特征的深度多任务学习框架（Deep Multi-task Learning framework with Relational Attention, DMLRA），给出连锁品牌的投资预测是一种可行的商业选址问题解决方案。基于该模型运行的结果，可以有效地为投资者提供若干最优的决策方案。本文也通过对文献中提及大量具体案例进行实验探究，验证了模型实施的可行性与有效性，以保证商业选址的推荐方法能在真正投入到生产实践中后能在最大程度上为投资者带来经济效益。

此外，本文通过研究发现，对于一个有效的商业选址推荐平台而言，需要综合考量时间与位置信息、消费者行为与心理学信息、舆情评论信息等。此外，还有一些学者，在选址平台中引入了建筑领域常用的建筑信息模型（BIM）的概念，可以深入商圈细节，从而达到进行辅助选址的作用，并能将体量庞大却联系较弱的海量数据从宏观角度综合演绎，无论是从经营策略、广告投放还是选址优化设计方面，都将为投资者提供敏锐的预测。在本文的最后，对采用多任务学习进行商业选址预测的工作进行总结，给出相关领域未来的发展前景预期，为相关研究提供辅助性的参考。

Keywords: 多任务学习, 商业选址推荐, 综述

1 引言

随着各种电子商务平台的日益兴盛，互联网每天都会产生大量的用户数据。这些丰富的用户数据资源与地理信息资源 POI，对商业选址都会提供强有力的数据支持。而目前大多数的商业选址研究更多地采用传统的机器学习与数据挖掘方法，来进行商业投资位置的选择。

然而，尽管近年来机器学习领域已经取得了突破性的进展，但对于预测商业选址的准确度方面，仍存在一定的技术瓶颈。首先，商业选址的成功判定标准涉及到多维度特征的提取，这是传统机器学习模型难以处理的 [14]。其次，不同的判定标准之间的相关性通常是复杂且无法预知的。此外，商店特征与投资成功率之间的映射关系对于传统方法而言也是十分棘手的问题。由此构成了我们将多任务学习模型应用到商业选址问题的研究动机。

传统的商业推荐算法还存在其他一些问题，如数据来源来自调查问卷，或店铺的历史运营数据；优化目标多以最大化店铺的覆盖面积作为目标；推荐方案更多地依赖于决策者经验和传统的运筹学方法，等等。

本文的目的在于，对文献中提及的数据组织和管理方法，以及多任务学习模型在商业选址问题的应用进行综述，使得推荐方案能够最大限度地，接近选址推荐的目标。由于商业选址是一类十分宽泛的问题，因而在本综述文章中，将选址推荐的目标聚焦于连锁店与连锁品牌的选址推荐。



图 1. 问题描述

连锁店，是指若干共用同样的品牌名称，有统一的经营管理和商业运营标准，并分布于不同地理位置的一些店铺的总称 [19]。近年来，随着社会经济的持续发展和企业的不断扩张，连锁店的经营模式得到了更为广泛的应用。如餐饮业的海底捞火锅店、麦当劳、星巴克，服饰业的 H&M、Nike、Zara 等品牌的迅猛发展，连锁店这种商业模式开始逐渐在市场中占据主导地位。由此为这些连锁品牌带来一个关乎企业发展的核心问题，即连锁店的选址问题。因此有必要对海量的用户产生数据进行分析，从而为投资者给出最优化的选址推荐。

2 研究现状

2.1 商业选址的研究方法

在商业选址方面，既有的研究成果主要利用了丰富的地理信息和社交媒体数据，对商业选址成功率给出具体预测结果。由于智能手机逐渐被大多数人使用，和 GPS、北斗导航定位技术的普及，商业选址的数据来源也得到了极大的丰富 [11]。

在一些文献中，作者通过对商店地理位置数据，使用一种基于位置的社交网络，成功地预测了商店每日顾客到达数 [16]；此外，有的学者提出了一种根据连锁店规模大小而进行的选址推荐框架，并给出正确引入这种经营模式的可行方案 [15]；也有研究人员通过百度地图的数据信息，对消费者需求和行为进行分析，给出最优化的选址方案 [22]；还有一些研究人员，将商业选址问题与时下快速发展的全球地理信息系统 GIS 相结合，生成选址的推荐方案 [13]。

表 1. 商业选址推荐方法

| 方法名称 | 数据来源 | 方法核心 |
|---------|------------|-------------------------|
| LBSN 选址 | 手机定位信息 | 利用用户使用移动手机产生的地理信息数据进行推荐 |
| 规模选址 | 店铺运营数据 | 根据店铺经营规模数据进行选址推荐 |
| 地图数据选址 | 百度地图数据 | 通过百度地图的用户查询数据进行选址推荐 |
| GIS 选址 | 全球地理信息系统数据 | 利用全球地理信息系统的辅助分析进行选址推荐 |

总体而言，数据驱动的选址方法，逐渐成为了解决此类问题的主流方法。然而现有的方法存在以下弱点：1) 这些方法对选址成功性的判定主要考量

单一标准，而事实上选址有效性的衡量是一个多维度评估问题；2）既有的预测模型主要应用线性模型或浅层模型，这样的预测模型不足以支持和处理多变的数据特征与预测结果之间的复杂联系。

2.2 多任务学习的研究现状

对商业选址的相关评判指标的分析过程，其实可以抽象表示成一个对相关评价指标的分层次分类过程。通过对相关指标的预测，生成推荐方案。以餐饮连锁店为例，通过预测在不同目标点处可能取得的用户评分指标以及用户评论数量，可以推测出该投资的目标餐饮店的受欢迎程度，和顾客访问量。

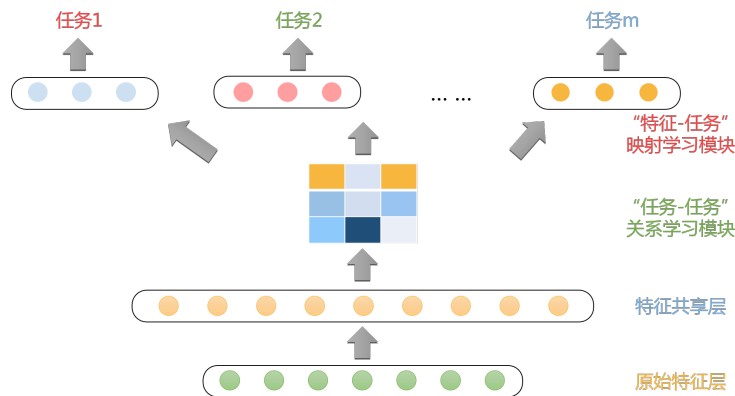


图 2. 多任务学习模型概述

目前对深度多任务学习算法的研究所提供的问题解决方案，可以概括地表示为图 2 所示的模型架构，这种结构化的算法可以更好地兼顾不同类别任务之间的层次化结构 [21]。与传统的深度卷积神经网络相比，这种深度多任务学习算法能够简化平衡分类器的训练，使得不同种类的评判指标信息可以更好的指导深度神经网络的训练 [23]。

通过使用这种多指标的联合评估方式，减小了模型不同层之间的训练损失传递，从而更好地提高分类别预测的精确度。

3 多任务学习模型核心理论

3.1 特征提取

有关商业选址评判的特征提取的理论研究有很多，从大量文献记述中可以总结为从城市维度、商业区维度、商场维度三个层次，来展开的特征提取。在城市维度，现有方法中主要考虑了商场所处的商业区的等级以及功能特性 [18]。在商业区维度上，则更多地考虑了商业区的商场密度、商业区对顾客的吸引力、商业区内同类型商店所带来的竞争指数以及人流量 (通过计算出租车、地铁等 dropoff 数据来估计) [24]。在最内部的商场维度上，有的研究人员则是从大众点评数据上提取了人均消费、商场大小、促销情况、商场品牌质量 (口碑评分) 等多种维度的特征 [20]。将提取的多维度、多种类的特征作为模型的输入进行训练。

此外，还有学者建议将上述指标整合成内部特征与外部特征两类指标进行评价。[17] 所谓内部特征，就是指商圈内部对消费者的作用与影响的特征，包括商圈 POI 数目、商场竞争性与互补性特征等等。而外部特征，则是指商业区周边的环境信息，包括商场周边的人流量数据、交通流量数据以及消费行为分析数据等等，一些典型的特征提取的分类可以用表 2所示。

表 2. Mood quadrants and their corresponding number of songs

| 种类 | 特征实例 |
|--------|-----------|
| 内部特征举例 | 店铺平均消费水平 |
| | 连锁品牌的资产净值 |
| | 商品种类丰富程度 |
| | 连锁店占地面积 |
| | 商圈吸引力指数 |
| | 商场竞争力指数 |
| 外部特征举例 | 行人流量密度 |
| | 居民区密集度 |
| | 地铁节点流量密度 |
| | 出租车卸客数 |
| | 易达性 |
| | 周边公交站数目 |

3.2 多任务学习模型

传统的多任务学习模型主要由 4 部分组成，包括特征编码层、任务层、关联表征的多任务层和预测层这四个部分 [1] [2]，我们可以从图 2 中看出一种深度多任务学习模型的典型架构图。我们将上一步提取出来的多维度多种类的特征首先需要送入编码层进行编码，对不同类型以及维度的输入特征向量进行线性变换后再通过 Relu 层，然后送入任务层进行交叉学习，最后通过预测层进行最终的结果预测 [12]。

特征编码层 在这一层中，主要完成对各目标商业区提供的源数据呈现共享表征进行提取，并形成特征向量编码，供下一层的使用 [6]。在本层所形成的编码可以表示成如下形式。对于由 n 个连锁店组成的集合 \mathbf{S} 中的每一家连锁店 $\mathbf{s}_i \in \mathbf{S}$ 而言，其内部特征可以表示为

$$\mathbf{x}_i^I = [x_{i,1}^I, x_{i,2}^I, x_{i,3}^I, \dots, x_{i,D^I}^I]^\top \in R^{D^I}$$

而外部特征则可以用

$$\mathbf{x}_i^E = [x_{i,1}^E, x_{i,2}^E, x_{i,3}^E, \dots, x_{i,D^E}^E]^\top \in R^{D^E}$$

来进行表示 [9]。因此，特征编码层的特征向量可以写作

$$\mathbf{x}_i = [\mathbf{x}_i^I; \mathbf{x}_i^E] \in \mathbf{R}^D$$

其中 D 表示内部与外部的特征空间 $D = D^I + D^E$ 。

任务层 在与任务层相关的研究中，一些研究人员将特征编码层得到的进行编码后特征记为 h_{ENC} ，用公式 1 描述特定任务层的数学模型

$$c^k = ReLU(W_{TS}^k h_{ENC} + b_{TS}^k) \quad (1)$$

其中 b_{TS}^k 是偏置层 (*bias*), c^k 是将原来的多维多种类特征进行进一步提取得到的降维后任务 k 的新特征 [3]。

关联表征的多任务层 为了得到不同任务之间的关系，Wang 等人的研究工作是在任务层之上引入了一个关系矩阵 R 来表示一个新的关联表征多任务

层 [8]。上层输出的 c^k 是第 k 个任务的特征，我们定义在交叉学习 k 个任务之后得到的新的隐藏层特征 $\alpha^1, \alpha^2, \alpha^3, \dots, \alpha^k$ 的数学模型如下

$$\begin{bmatrix} \alpha^1 \\ \vdots \\ \alpha^k \end{bmatrix} = \begin{bmatrix} R_{11} & \cdots & R_{1K} \\ \vdots & \ddots & \vdots \\ R_{K1} & \cdots & R_{KK} \end{bmatrix} \cdot \begin{bmatrix} c^1 \\ \vdots \\ c^k \end{bmatrix}$$

在该层的最后，Xu 等人通过采用池化（MaxPooling）操作和 sigmoid 函数操作，将多个任务之间重要的信息特征再次提取出来，得到新的特征 $\tilde{c}^1 \dots \tilde{c}^k$ [10]。

预测层 预测层是整个多任务学习模型的顶层输出，也是多任务学习的推荐目标。预测层的核心是一个线性变换层，输出 y^k 为相应的预测结果 [5]。

$$y^k = w^k \tilde{c}^k + b^k$$

损失函数 为了评估模型的预测效果，并对模型进行调整与优化，有必要引入一个合适的损失函数。Jing 等人在其相关研究中，考虑使用最小平方损失 (least-square loss function) 作为多任务学习模型的损失函数进行评估 [4]，针对这一类选址问题，这个损失函数可以用公式 2 表示

$$Loss = \sum_{i=1}^k \sum_{k=1}^K \|y_{ireal}^k - y_i^k\|^2 + \mu \|w\|_F^2 + \beta \|R\|_F^2 + Loss_{ENC} \quad (2)$$

损失函数由预测值和真实值的 $L2$ 范数、任务层的参数 W 、多任务之间的关系矩阵 R 以及编码过程产生的 Loss 这 4 部分组成。在 Niu 等人的书籍文献中，提到了使用 Vizier 去寻找最优参数的方法，优化过程使用加速梯度下降方法 AGM(accelerate gradient method)，也都产生了比较好的预测效果 [7]。

3.3 实验结果

其次，我们可以用这些商场的内部特征与外部特征信息，对上述预测结果进行更深入的理解。通过出租车接送客数据信息，与公共交通自助交易信息可知，鼓楼附近的交通情况十分便捷。此外，同类型的店铺数量较少，结合上述两点分析，可以带来较多的客流量。但由于缺乏竞争，就可能导致该店铺的经营与管理不利，致使其服务水平下降，影响用户的服务评分。综上所述，案例探究得出的结论与选址推荐系统预测值基本吻合。

4 结论

其次, 我们可以用这些商场的内部特征与外部特征信息, 对上述预测结果进行更深入的理解。通过出租车接送客数据信息, 与公共交通自助交易信息可知, 鼓楼附近的交通情况十分便捷。此外, 同类型的店铺数量较少, 结合上述两点分析, 可以带来较多的客流量。但由于缺乏竞争, 就可能导致该店铺的经营与管理不利, 致使其服务水平下降, 影响用户的服务评分。综上所述, 案例探究得出的结论与选址推荐系统预测值基本吻合。

Acknowledgements This work was partially supported by the National Natural Science Foundation of China (No. 61332018), the National Department Public Benefit Research Foundation (No. 201510209), and the Fundamental Research Funds for the Central Universities.

参考文献

1. Caruana, R.: Multitask learning. *Machine Learning* 28(1), 41–75 (1997)
2. Caruana, R.: A dozen tricks with multitask learning (2012)
3. Held, M., Rabe, A., Senf, C., Linden, S.V.D., Hostert, P.: Analyzing hyperspectral and hypertemporal data by decoupling feature redundancy and feature relevance. *IEEE Geoscience and Remote Sensing Letters* 12(5), 983–987 (2017)
4. Jing, L., Guo, B., Zhu, W., Li, M., Yu, Z.: Where to place the next outlet? harnessing cross-space urban data for multi-scale chain store recommendation. In: *Acm International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct* (2016)
5. Long, M., Cao, Z., Wang, J., Yu, P.S.: Learning multiple tasks with multilinear relationship networks. *Computer Science* (2017)
6. Mariani, M., Baggio, R., Fuchs, M., Höpken, W.: Business intelligence and big data in hospitality and tourism: a systematic literature review. *International Journal of Contemporary Hospitality Management* 30(12), 3514–3554 (2018)
7. Niu, H., Liu, J., Fu, Y., Liu, Y., Bo, L.: Exploiting Human Mobility Patterns for Gas Station Site Selection (2016)
8. Wang, F., Chen, L., Pan, W.: Where to place your next restaurant?: Optimal restaurant placement via leveraging user-generated reviews. In: *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. pp. 2371–2376. ACM (2016)

9. Wang, Q., He, M., Du, M., Chow, S.S.M., Lai, R.W.F., Zou, Q.: Searchable encryption over feature-rich data. *IEEE Transactions on Dependable and Secure Computing* PP(99), 1–1 (2018)
10. Xu, M., Wang, T., Wu, Z., Zhou, J., Li, J., Wu, H.: Demand driven store site selection via multiple spatial-temporal data. In: *Proceedings of the 24th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. p. 40. ACM (2016)
11. Zhao, S., King, I., Lyu, M.R., Zeng, J., Yuan, M.: Mining business opportunities from location-based social networks. In: *Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*. pp. 1037–1040. ACM (2017)
12. Zheng, Y., Capra, L., Wolfson, O., Yang, H.: Urban computing: concepts, methodologies, and applications. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5(3), 38 (2014)
13. 何洁昕: 基于 GIS 的商业选址系统的设计与实现. Master's thesis, 河北科技大学 (2018)
14. 周晓东, 刘娜, 蔡文: 基于 qfd 的商业银行网点选址应用研究. *经营管理者* 4, 2–3 (2015)
15. 安芸睿, et al.: 中国城市中外国连锁店的发展探析. *山西建筑* (2018 年 20), 23–25 (2018)
16. 屈弘扬, 於志文, 田苗, 郭斌: 基于 lbsn 的商业选址推荐系统的研究与实现. *计算机科学* 42(9), 33–36 (2015)
17. 毕璋: 基于 POI 数据的大型铁路客运站站区商业空间分布特点研究. Ph.D. thesis, 西南交通大学 (2018)
18. 王晓梦, 王锦, 朱青: 基于签到数据的城市商业空间空心化识别研究——以北京市城六区为例. *城市发展研究* 25(2) (2018)
19. 董鸿安, 丁镭, 林远贵: 基于 sems 的本土咖啡连锁店体验营销策略研究. *商业经济* (4), 79–82 (2018)
20. 薛冰, 肖骁, 李京忠, 姜璐, 谢潇: 基于 poi 大数据的城市零售业空间热点分析——以辽宁省沈阳市为例. *经济地理* 38(5), 36–43 (2018)
21. 赵其鲁, 李宗民: 基于深度多任务学习的层次分类. *计算机辅助设计与图形学学报* 30(5), 886–892 (2018)
22. 邓悦: 基于多源兴趣点数据的城市功能区划分方法研究. Master's thesis, 中国测绘科学研究院 (2018)
23. 马建阳, 张宝鹏: 基于多任务学习的多源数据分类研究. *计算机应用研究* 35(11), 34–37 (2018)

24. 高忠文, 牛孜飏: 基于大数据的交通流量模式分析. 哈尔滨理工大学学报 23(06), 124–127 (2018)