

# 基于多任务学习的商业选址推荐方法的研究

匿名

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。并且对多任务机器学习方法在商业领域的应用情况进行了综述。本文运用研究人员在公开数据集上的实验数据，对文献中提到的方法的效果展开了实验评估。实验数据表明，目前在深度多任务学习领域提出的新方法，即通过提取出与商场有关的内部与外部特征，并应用关联特征的深度多任务学习模型（Deep Multi-task Learning framework with Relational Attention, DMLRA），相较于其他方法，提供了一种更为可靠的商业选址问题解决方案。基于该模型的预测结果，可以有效地为投资者提供若干最优的决策方案。本文也对模型实施的可行性与有效性进行了评估，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，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方法而言，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。选址过程中需要深入商圈细节，从而达到进行辅助选址的作用，并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，为相关研究提供参考。

**Keywords:** 多任务学习, 商业选址推荐, 文献综述

## 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。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，但对于提高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 [36]。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。由此越来越多的学者开始考虑将多任务学习模型应用到商业选址当中。

传统的商业推荐算法还存在其他一些问题，如数据源来自调查问卷，或店铺的历史运营数据等等，导致大量用户产生数据没能得到有效利用；优化目标多以最大化店铺的覆盖面积作为目标，不能兼顾到影响商业投资的多方面因素 [16]；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

随着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，即下一家连锁店开在什么位置，最利于新店的发展。因此有必要对海量的用户数据进行分析，从而为投资者提供最优的选址推荐 [44]。

## 2 综述模型

本文的综述模型如图 2所示。

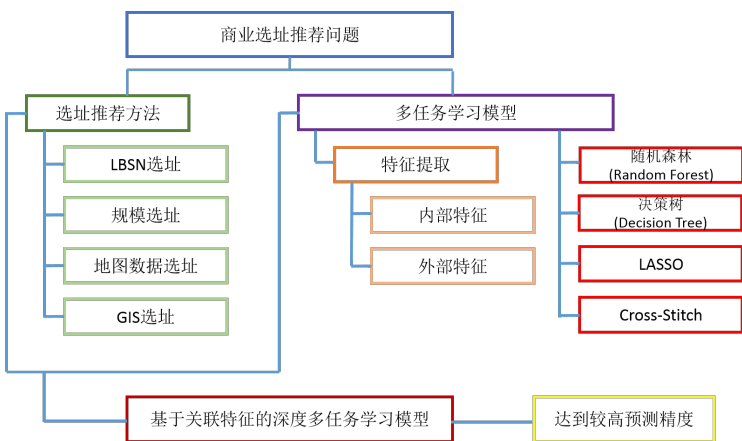


图 2. 综述模型

## 3 研究现状

### 3.1 商业选址的传统推荐方法

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 [4, 33]。

在屈弘扬等人的研究中，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 [39]；此外，还有一些

研究人员分别提出了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 [8,38]；在邓悦的文章中，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 [48]；在何洁昕的实验方法中，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 [35]。

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。

3.2 多任务学习的研究现状

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 3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 [47]。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 [49]。

本文认为，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。

3.3 二者之间的联系

无论是单独进行选址预测，还是单独运用多任务机器学习模型，在选址预测上，都将有失偏颇。现今有不少学者，将二者综合运用起来，将传统商业选址预测问题中考虑的特征引入到多任务学习模型当中。商业选址的相关

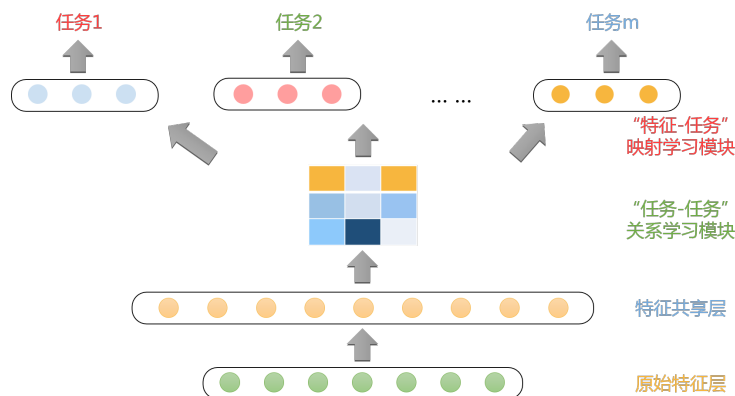


图 3. 多任务学习模型结构示意图

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，和顾客访问量等等。

## 4 多任务学习模型核心理论

### 4.1 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 [42]。在商业区维度上，则更多地考虑了商业区的商店密度、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行人密度等。高忠文等人通过计算出租车、地铁等上下乘客数据来估计人流量，从中提取特征 [50]。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 [46]。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 [41]。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. 内部与外部特征举例

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

#### 4.2 多任务学习模型

传统的多任务学习模型主要由四部分组成，包括特征编码层、任务层、关联表征的多任务层和预测层 [1,2]，我们可以从图 3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 [10,34]。

**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 [20]。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 [30]。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 [12]。

**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 [29]。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  $\tilde{c}^1 \dots \tilde{c}^k$  [14, 31]。

**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 [19]。

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

**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 [15]，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 Golovin 等人的

实验中，提到了使用 Vizier 去寻找最优参数的方法 [9]，而在 Niu 等人的书籍文献中，在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 [22]。

### 4.3 城市商业区划分

商业区的等级与功能划分，也是商业选址问题中，需要重点考量的因素。同时作为深度多任务学习模型输入的一个维度的评价指标，将为模型的训练提供参考，因此对城市商业区分布的研究便显得尤为重要。这些年来，地图数据的逐步开源，逐渐带来了相应的背景数据的开源（又称兴趣点 POI）[5, 32]。越来越多的学者想要通过对这些数据的特征来提出一种全新的方法识别商业区以及分类 [43]。

**商业区的识别** 要完成对商业区的划分，首先需要通过一定的指标将商业区的范围进行圈定。参考人们商业活动的习惯，一些研究人员发现，人们的商业活动的基本单位是街区。因而在城市商业区划分问题中，可以选定街区作为基本的划分单位。例如 Celikoglu 等人将城市按照不同类型的路网分割、进行聚类计算 [3]。

**城市商业等级划分** 按照商业活动量的差异，相关研究采用了自然断裂点分类法划分城市的所有商业区为三个等级。划分的主要依据是商业活动量，根据活动量的频率分布直方图，将频率分布突变点作为划分区间端点，划分商业区等级 [40]。

**城市商业功能区划分** 在识别好商业区之后，可以按照一定的指标，参考 K-Means 聚类算法的相关实现原理，将所有的商业区划分为饮食文化型、专营型、购物中心型、便利型、综合型商业区这五种类型 [7]。在这五种类型中，举例来说，饮食文化型商业区聚集了各大餐饮连锁。便利型商业区主要是为周边小区提供生活必需品以及基本生活服务。专营型商业区则是某一类型的商品的贩卖聚集地，常见的有家电电子市场，建材市场等。购物中心型商业区服务类型丰富，服务辐射范围较广。综合型商业区主要包括综合市场、各类专卖店的聚集地等。



## 5 实验探究

针对商业选址这类问题，已有不少学者从多种角度开展了相关实验。以下汇总了使用不同方法，对选址评价指标的预测准确度进行了总结。

### 5.1 实验数据准备

一些学者为了评估关联表征的多任务学习模型在预测精确度上相较于其他方法的优越性，分别采集了以下三个来自北京市的真实数据集进行验证。大众点评数据：这一数据集主要描述了商业区功能、评级以及评分信息，包含大众点评网上共计 456 个商场 10070 家餐饮店以及 6028 家服饰连锁店的信息 [37]。城市交通 AFC 自助交易记录：在这一实验中，研究人员采集了城市地铁和公交线路乘客的购票记录，采集区间为 2017 年 6 月 5 日至 6 月 17 日，共计 114232304 条数据记录 [51]。行车轨迹记录：根据订车平台上以及出租车 GPS 的行车轨迹记录的数据信息，提取商场周边的人流量信息 [45]。

### 5.2 评价指标

为了给出准确性预测，实验随机地将数据集划分训练集和测试集，其中训练集占总数据集规模的 60%，测试集占 40%。对于不同选址模型的评估指标为方均根误差（RMSE）[6]。和皮尔逊相关系数（PCC）。其中，PCC 是衡量两个随机变量之间的线性相关性的指标 [25]，RMSE 指标被广泛的应用于回归误差的分析，其定义如下

$$RMSE = \sqrt{\frac{\sum_{i=1}^k \sum_{k=1}^K (y_i^k - \hat{y}_i^k)^2}{nK}}$$

### 5.3 实验应用方法

实验中所采用的方法如表 3 中所示。

### 5.4 实验结果分析

表 4 和表 5 的实验结果，分别展现了采用关联特征的深度多任务学习模型（Deep Multi-task Learning framework with Relational Attention, DMLRA），和其他传统模型，在餐饮和服饰两个数据集上的预测结果的 RMSE 指标。

表 3. 实验方法列表

方法名称	方法描述
RF	随机森林 (Random Forest) 是一种利用多棵树对样本进行训练和预测的一种分类器，实验中选择的随机森林的树的数量为 60 棵 [24]。
DT	决策树 (Decision Tree) 是一种常用的树形结构分类方法在实验中约定树的最大深度为 2 [26]。
KNN_unif	这种方法通过联合权重训练一个回归模型在实验中为 KNN 网络使用 6 个最近邻的邻居节点 [11]。
GBR	梯度加速回归算法 (Gradient Boosted Regression) 设置 50 棵回归树，每棵树的深度为 2 [27]。
LASSO	这是一种单任务学习方法每个任务之间彼此独立采用 L1 范数正则化进行计算 [23]。
Adaboost	这是一种适应性加速算法，实验中配置一棵“弱学习树”并设置集成学习周期为 50 [17]。
SVR_rbf	一种基于支持向量机的回归算法，该方法使用径向基函数内核，并设置 $\gamma = 0.01$ [18]。
ANN	一种简单的单层隐藏层的神经网络。
Calibration	标准化多任务学习方法 (Calibration) 加入了一种基于典型的多任务学习标准正则化矩阵 [13]。
Shared-Bottom	一种使用共享隐藏层的深度多任务学习方法。
L2-Constrained	一种使用 L2 距离进行共享层正则化的深度多任务学习方法 [28]。
Cross-Stitch	一种通过 Cross-stitch 单元来共享两个任务之间信息的深度多任务学习方法 [21]。

表 4. 基于餐饮连锁店数据集的各种预测方法的 RMSE 指标

	y1	y2	y3	y4	y4
RF	1156.625	3.921	0.518	0.521	0.567
DT	1278.090	4.109	0.543	0.581	0.604
KNN_unif	1364.145	4.401	0.594	0.597	0.620
GBR	1317.825	4.272	0.571	0.601	0.608
LASSO	1171.059	3.922	0.595	0.625	0.646
Adaboost	1404.083	4.198	0.575	0.598	0.600
SVR_rbf	1203.038	4.337	0.573	0.565	0.603
ANN	1014.505	3.784	0.501	0.560	0.547
Calibration	1226.450	3.767	0.499	0.502	0.544
Shared-Bottom	1065.990	3.932	0.546	0.544	0.589
L2-Constrained	958.086	3.829	0.572	0.538	0.565
Cross-Stitch	1068.238	3.744	0.533	0.494	0.540
<b>DMLRA</b>	<b>933.406</b>	<b>3.611</b>	<b>0.488</b>	<b>0.479</b>	<b>0.536</b>

首先，随机森林、决策树等传统的单一任务学习方法和 cross-stitch 这种多任务学习方法更多的依赖于任务关系假设，这就导致了这些方法相比于深度多任务学习模型而言灵活度更低。观测数据显示，DMLRA 可以更有效地实现正负样本转换控制的自适应。

表 5. 基于服饰连锁品牌数据集的各种预测方法的 RMSE 指标

	y1	y2	y3	y4	y4
RF	33.555	3.149	0.439	0.439	0.459
DT	53.900	3.319	0.475	0.477	0.468
KNN_unif	53.580	3.517	0.0.491	0.493	0.494
GBR	51.088	3.111	0.451	0.463	0.457
LASSO	44.849	3.101	0.482	0.487	0.469
Adaboost	58.512	3.081	0.462	0.449	0.448
SVR_rbf	53.734	3.250	0.510	0.468	0.504
ANN	36.083	3.259	0.477	0.438	0.443
Calibration	43.099	2.934	0.432	0.431	0.444
Shared-Bottom	36.528	2.973	0.467	0.474	0.455
L2-Constrained	30.011	2.946	0.466	0.460	0.466
Cross-Stitch	29.845	2.818	0.399	0.404	0.416
<b>DMLRA</b>	<b>23.287</b>	<b>2.668</b>	<b>0.361</b>	<b>0.386</b>	<b>0.397</b>

其次, 可以从实验结果中注意到, 采用了 DMLRA 模型的方法进行预测可以在大多数任务上, 取得显著更优的 RMSE 和 PCC 指标。

## 6 结论

本文综述了为连锁品牌新店选址进行推荐的方法。首先从商圈的特征提取入手, 将与选址评判相关的因素, 整合成内部与外部两类特征, 作为深度多任务学习模型的输入。其次, 在论述传统的多任务学习方法的基础上, 目前新提出的基于关联特征的深度多任务学习 (DMLRA) 方法, 能够自适应地控制正负样本之间的转换。特别地, 该方法引入了一种特殊任务追踪机制来为不同的学习任务指定特征, 因为商业选址成功性判定的准则依赖于共享特征集上的不同维度。另外, 该方法还设计了一个关联学习模型来区分多重准则之间的联系。为了评估模型的效果, 本文引用了部分学者在真实数据集的基础上, 开展的大量实验, 并分别与传统方法和目前最先进的方法进行验证, 以给出性能指标评价结果。实验结果显示, 基于关联特征的深度多任务学习模型的实现效果, 在一定程度上优于传统方法的实现效果。

因此, 在未来的商业选址预测任务中, 考虑多指标评估的关联特征深度多任务学习方法, 将极有可能会成为主流的研究方向, 并有希望与大众点评、美团等电子商务平台展开合作, 将该模型投入到实际的生产应用当中。

**致谢** 首先要感谢实验室的老师, 对我完成本篇文献综述过程中所提供的指导。此外还要感谢荣文戈老师以及我身边的同学, 对我在使用文本编辑工具方面提供的帮助。最后要感谢家人和朋友对我一直以来的支持与关爱。

## 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. Celikoglu, H.B., Silgu, M.A.: Extension of traffic flow pattern dynamic classification by a macroscopic model using multivariate clustering. *Transportation Science* 50(3), 966–981 (2016)
4. Chen, T.Y., Chen, L.C., Chen, Y.M.: Mining location-based service data for feature construction in retail store recommendation. In: *Industrial Conference on Data Mining*. pp. 68–77. Springer (2017)

5. Feng, S., Li, X., Zeng, Y., Cong, G., Chee, Y.M., Yuan, Q.: Personalized ranking metric embedding for next new poi recommendation. In: Twenty-Fourth International Joint Conference on Artificial Intelligence (2015)
6. Fortin, V., Abaza, M., Anctil, F., Turcotte, R.: Why should ensemble spread match the rmse of the ensemble mean?(vol 15, pg 1708, 2014). JOURNAL OF HYDROMETEOROLOGY 16(1), 484–484 (2015)
7. Frandsen, P.B., Calcott, B., Mayer, C., Lanfear, R.: Automatic selection of partitioning schemes for phylogenetic analyses using iterative k-means clustering of site rates. BMC evolutionary biology 15(1), 13 (2015)
8. Fu, Y., Hui, X.: Modeling of geographic dependencies for real estate ranking on site selection. In: IEEE International Conference on Data Mining Workshop (2015)
9. Golovin, D., Solnik, B., Moitra, S., Kochanski, G., Sculley, D.: Google vizier: A service for black-box optimization. In: Acm Sigkdd International Conference on Knowledge Discovery and Data Mining (2017)
10. Gong, P., Zhou, J., Wei, F., Ye, J.: Efficient multi-task feature learning with calibration. In: Acm Sigkdd International Conference on Knowledge Discovery and Data Mining (2014)
11. Guo, G., Wang, H., Bell, D., Bi, Y., Greer, K.: Knn model-based approach in classification. In: OTM Confederated International Conferences” On the Move to Meaningful Internet Systems”. pp. 986–996. Springer (2003)
12. 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)
13. Hirata, C., Seljak, U.: Shear calibration biases in weak-lensing surveys. Monthly Notices of the Royal Astronomical Society 343(2), 459–480 (2018)
14. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
15. 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)
16. Jou, B., Chang, S.F.: Deep cross residual learning for multitask visual recognition. In: Proceedings of the 24th ACM international conference on Multimedia. pp. 998–1007. ACM (2016)
17. Li, X., Wang, L., Sung, E.: Adaboost with svm-based component classifiers. Engineering Applications of Artificial Intelligence 21(5), 785–795 (2008)

18. Liu, L., Lei, Y.: An accurate ecological footprint analysis and prediction for beijing based on svm model. *Ecological Informatics* 44, S1574954117302972 (2018)
19. Long, M., Cao, Z., Wang, J., Yu, P.S.: Learning multiple tasks with multilinear relationship networks. *Computer Science* (2017)
20. Mariani, M., Baggio, R., Fuchs, M., Höepken, 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)
21. Misra, I., Shrivastava, A., Gupta, A., Hebert, M.: Cross-stitch networks for multi-task learning. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 3994–4003 (2016)
22. Niu, H., Liu, J., Fu, Y., Liu, Y., Bo, L.: Exploiting Human Mobility Patterns for Gas Station Site Selection (2016)
23. Park, T., Casella, G.: The bayesian lasso. *Publications of the American Statistical Association* 103(482), 681–686 (2008)
24. Paul, A., Mukherjee, D.P., Das, P., Gangopadhyay, A., Chintla, A.R., Kundu, S.: Improved random forest for classification. *IEEE Transactions on Image Processing* 27(8), 4012–4024 (2018)
25. Schober, P., Boer, C., Schwarte, L.A.: Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia* 126(5), 1763–1768 (2018)
26. Siu, C.: Automatic induction of neural network decision tree algorithms (2018)
27. Tyree, S., Weinberger, K.Q., Agrawal, K., Paykin, J.: Parallel boosted regression trees for web search ranking. In: *International Conference on World Wide Web* (2011)
28. Vilbe, P., Calvez, L.C.: Constrained l2 suboptimal model reduction. *Electronics Letters* 23(25), 1340–1342 (2007)
29. 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)
30. 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)
31. 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)

32. Yin, H., Zhou, X., Cui, B., Wang, H., Zheng, K., Nguyen, Q.V.H.: Adapting to user interest drift for poi recommendation. *IEEE Transactions on Knowledge and Data Engineering* 28(10), 2566–2581 (2016)
33. 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)
34. 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)
35. 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)
36. 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)
37. 夏令军, 刘艳芳, 刘国炜: 中国地级城市餐饮业分布格局及影响因素——基于“大众点评网”数据的实证研究. *经济地理* (5), 17 (2018)
38. 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)
39. 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)
40. 李鹏, 王英杰, 虞虎, 武文波, 马楠: 基于 gis 格网化分析支撑的旅游空间规划技术方法研究——以青岛市为例. *自然资源学报* 33(5), 813–827 (2018)
41. 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)
42. 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)
43. 王芳, 高晓路, 许泽宁: 基于街区尺度的城市商业区识别与分类及其空间分布格局——以北京为例. *地理研究* 34(6), 1125–1134 (2015)
44. 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)
45. 蔡宁, 贺锦江, 王节祥: “互联网 +”背景下的制度压力与企业创业战略选择——基于滴滴出行平台的案例研究. *中国工业经济* (3), 174–192 (2017)
46. 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)
47. 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)

48. 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)
49. 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)
50. 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)
51. 黄洁, 王姣娥, 靳海涛, et al.: 北京市地铁客流的时空分布格局及特征——基于智能交通卡数据. 地理科学进展 37(3), 397–406 (2018)