

基于居民感知视角的城市大脑建设方向分析

——以杭州市为例

刘巧巧¹, 夏汀舟², 黄贻强², 李涵圳³, 张文婧⁴

¹杭州电子科技大学经济学院, 浙江 杭州

²杭州电子科技大学卓越学院, 浙江 杭州

³杭州电子科技大学管理学院, 浙江 杭州

⁴杭州电子科技大学人文艺术与数字媒体学院, 浙江 杭州

收稿日期: 2022年9月18日; 录用日期: 2022年10月8日; 发布日期: 2022年10月20日

摘要

本文以杭州市居民为调查对象, 旨在了解杭州城市大脑建设现状及居民满意度, 并根据调查结果提出合理建议。采用分层抽样与偶遇抽样相结合进行问卷调查, 通过构建随机森林模型分析居民满意度的影响因子。研究表明, 不同年龄层群体对城市大脑了解存在差异性, 卫生健康和便民服务模块使用人数最多, 特定人群对特定模块有对应需求, 总体满意度受便民服务、交通出行等模块的影响较大。因此, 建议主管部门应制定科学合理规划, 形成有力技术支撑, 基层社区应加强对内宣传力度, 建设完善社区微脑, 居民应配合城市大脑建设, 积极发表反馈意见。

关键词

居民感知, 城市大脑建设, 随机森林

Direction Analysis of Urban Brain Construction Based on Residents' Perception Perspective

—Taking Hangzhou as an Example

Qiaoqiao Liu¹, Tingzhou Xia², Yiqiang Huang², Hanzhen Li³, Wenjing Zhang⁴

¹Department of Economics, Hangzhou Dianzi University, Hangzhou Zhejiang

²Department of Honors, Hangzhou Dianzi University, Hangzhou Zhejiang

³Department of Management, Hangzhou Dianzi University, Hangzhou Zhejiang

⁴Department of Media and Design, Hangzhou Dianzi University, Hangzhou Zhejiang

Abstract

This article takes Hangzhou residents as the survey object, aiming to understand the current situation of Hangzhou urban brain construction and residents' satisfaction, and put forward reasonable suggestions according to the survey results. Stratified sampling and encounter sampling were used to conduct a questionnaire survey, and a random forest model was constructed to analyze the influencing factors of residents' satisfaction. Studies show that different age groups have different understanding of urban brain, and the number of users of health and convenience service modules is the largest. Specific groups have corresponding needs for specific modules, and the overall satisfaction is greatly affected by convenience services, transportation and other modules. Therefore, it is suggested that the competent departments should make scientific and reasonable planning to form strong technical support, grassroots communities should strengthen internal publicity, build and improve community microbrain, and residents should cooperate with the urban brain construction and actively express feedback.

Keywords

Resident Perception, Urban Brain Building, Random Forest

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

党的十九届六中全会提出“在社会建设上，大幅度提升社会治理社会化、法治化、智能化、专业化水平”的目标，为城市治理体系发展和转型升级提供重要参考方向。2020年3月31日，习近平总书记来到杭州城市大脑运营指挥中心，就杭州运用城市大脑推进城市治理体系和治理能力现代化进行考察调研。2021年国家“十四五”规划纲要中“加快数字化发展，建设数字中国”独立成篇，指出“分类分级推进新型智慧城市建设，推进城市数据大脑建设，探索建设数字孪生城市”。以此为背景，全国“城市大脑”建设开展得如火如荼[1]。本文基于居民感知视角，探索杭州城市大脑建设现状及居民满意度，为城市治理数字化转型和城市大脑建构与升级提供数据支撑和智力支持。

2. 研究设计

2.1. 问卷设计

在预调查、相关文献资料归纳的基础上，设计“杭州城市大脑建设现状及居民满意度”的调查问卷，调查问卷主要包括三部分。第一部分是调查者基本情况，包括性别、年龄、学历、职业和在杭居住时间。第二部分是调查者对杭州城市大脑建设现状的感知情况，包括了解程度、了解途径、已知模块、使用模块、使用频率、意见和建议等。第三部分是居民满意度，主要采用李克特五分评价标准对城市大脑各项模块进行评价，其中，“1 = 非常不满意”到“5 = 非常满意”[2]。

2.2. 数据采集

本研究的调查对象选取杭州市钱塘区、西湖区和余杭区居民,于2021年7月17日至7月29日在上述区域开展问卷调查,调查过程中实际发放问卷435份,回收428份,剔除回答不全、填写不完整及存在明显逻辑错误的问卷,最终得到有效问卷416份。总样本回收率98.39%,有效率97.20%。调查过程中采用一对一问卷调查方式。在被调查者提出疑问时,项目组成员耐心地进行解释,保证问卷填写质量及问卷整体有效率。

3. 研究结果

3.1. 样本结构分析

性别结构上,调查对象中男性占52.2%,女性占47.8%,性别比大致为1:1,具有一定合理性。年龄结构上,调查对象的年龄集中在18~44岁,这一群体对杭州城市大脑的了解程度和使用频率相对较高。职业结构上,以企业员工和学生为主,职业结构多样化。文化程度上,学历在专科、本科及以上的占比较大。在杭居住时间以1~5年和5~10年为主。

3.2. 建设现状

3.2.1. 了解现状

在了解程度方面,对城市大脑非常了解的共有63人,占比15.14%;基本了解的共有334人,占比80.29%;不了解的共有19人,占比4.57%。在了解途径方面,89%的被调查者通过网络了解杭州城市大脑,50%~60%的被调查者通过与他人交谈、电视新闻或社区宣传等方式了解杭州城市大脑。了解途径以网络为主,社区宣传力度还需加强。已知模块包括便民服务、卫生健康、交通出行、房屋管理和文化旅游[3]。其中便民服务、卫生健康和交通出行模块的已知人数占比均超过80%,是已知人数最多的三个模块。该结果反映出便民服务、卫生健康和交通出行服务已经深入市民日常生活,是市民对杭州城市大脑印象的重要来源。

3.2.2. 使用现状

在使用模块方面,调查对象使用最多的两个模块是卫生健康和便民服务,分别有95.0%和93.7%的被调查选择此两项。其次是交通出行,占比为65.7%。该结果反映出卫生健康、便民服务和交通出行模块在市民中有较高使用度。在使用频率方面,58.7%的被调查者使用城市大脑服务的频率为每周1~3次,23.7%的被调查者使用城市大脑服务的频率为每周4~5次。另外,有6.3%的被调查者的使用频率高于每周5次,有11.3%的被调查者使用频率低于每周1次。在改善模块方面,交通出行模块中认为需要改善的人数占使用人数的百分比最高,占比为62.84%;便民服务模块职工认为需要改善的人数占使用人数的百分比最低,占比为17.47%。

3.3. 居民满意度

3.3.1. 描述性统计

各类服务满意度平均分对比情况见图1。杭州城市大脑的总体满意度评分均值为4.24,且各项满意度均值全部大于4.00。这表明居民对杭州城市大脑总体满意。其中评价最高的是便民服务中的市民卡和证件办理,均值为4.72,表明居民认可这两个模块。评价最低的是交通出行中的交通信息查询,均值为4.20。有被调查者反映交通出行模块存在实时服务信息提供不精确和地理信息更新不及时的问题。



Figure 1. Average satisfaction score radar chart
图 1. 满意度平均分雷达图

3.3.2. 随机森林模型

随机森林是以决策树为基本分类器的一个集成学习模型，它包含多个由集成学习技术训练得到的决策树，当输入待分类的样本时，最终的分类结果由单个决策树的输出结果投票决定。随机森林克服了决策树过拟合问题，对噪声和异常值有较好的容忍性，对高维数据分类问题具有良好的可扩展性和并行性。此外，随机森林是由数据驱动的一种非参数分类方法，只需通过对给定样本的学习训练分类规则，并不需要分类的先验知识[4]。随机森林算法框架见图 2。

随机森林可以用于分类和回归。当因变量 Y 是离散变量时，是分类；当因变量 Y 是连续变量时，是回归。运用随机森林处理分类问题时，对于测试样本，森林中每棵决策树会给出最终类别，最后综合考虑森林内每一棵决策树的输出类别，以投票方式来决定测试样本的类别。处理回归问题时，则以每棵决策树输出的均值为最终结果[5]。

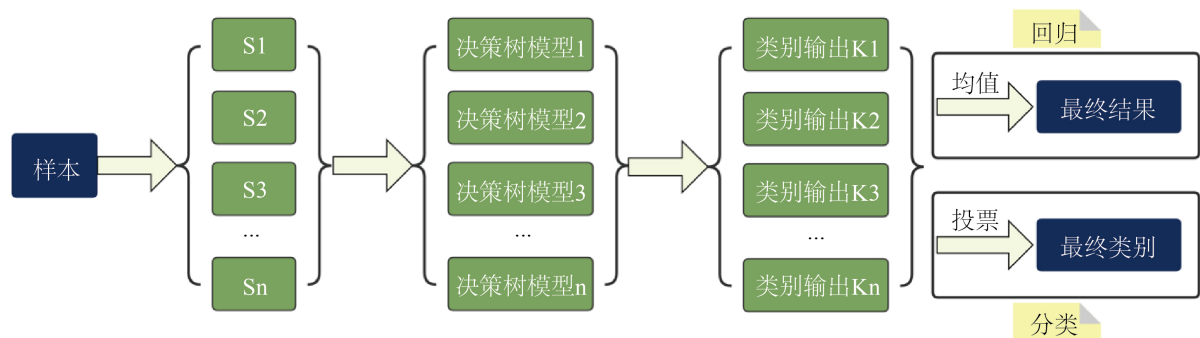


Figure 2. Random forest algorithm framework diagram
图 2. 随机森林算法框架图

具体操作思路为：

- 1) 从训练集中随机抽取一定数量的样本，作为每棵树的根节点样本；

2) 在建立决策树时, 随机抽取一定数量的候选属性, 从中选择最合适属性作为分裂节点;

3) 建立好随机森林以后, 对于测试样本, 进入每一颗决策树进行类型输出或回归输出; 若是分类问题, 以投票的方式输出最终类别; 是回归问题, 每一颗决策树输出的均值作为最终结果。

对总体满意度进行随机森林分析时, 由于我们把满意度认定在 1~5 的连续范围, 所以对连续值分析时采用随机森林回归分析。随机森林回归同时生成多个回归树做预测模型, 并将模型的结果汇总取平均值以提升预测模型预测的准确率。

对满意度模块中“市民卡”、“生活缴费”、“证件办理”、“健康码”、“舒心就医”、“医保查询”、“停车先离场”、“车位查询”、“交通信息查询”、“公积金”、“在线看房”、“房屋维修”、“旅游专线”、“20 秒景点入园”和“30 秒酒店入住”共 15 个子满意度模块进行分析, 由于满意度被认定为 1~5 的连续范围, 所以我们把这 15 模块进行定量分析, 连同总体满意度进行随机森林回归分析。将总体满意度作为训练集标签, 随机抽取样本的 25%作为训练集。

森林中树的数量对随机森林的性能影响至关重要, 为了获得更加客观, 拟合程度更高的最终结果, 我们选择将决策树数量定为 1000 棵。对满意度模块中的数据进行处理, 处理结果见表 1。

Table 1. Input data parameter table

表 1. 输入数据参数表

字段名	计数	平均数	标准差	最小值	1/4 分位数	1/2 分位数	3/4 分位数	最大值
市民卡	397	4.4257	1.2584	0	4.0000	5.0000	5.0000	5
生活缴费	397	4.3325	1.2654	0	4.0000	5.0000	5.0000	5
证件办理	397	4.4232	1.2481	0	4.0000	5.0000	5.0000	5
健康码	397	4.4030	1.1542	0	4.0000	5.0000	5.0000	5
舒心就医	397	4.1662	1.2155	0	4.0000	5.0000	5.0000	5
医保查询	397	4.2292	1.1914	0	4.0000	5.0000	5.0000	5
停车先离场后付费	397	2.8035	2.1583	0	0.0000	4.0000	5.0000	5
车位查询	397	2.8136	2.1732	0	0.0000	4.0000	5.0000	5
交通信息查询	397	2.7204	2.1106	0	0.0000	3.0000	5.0000	5
公积金	397	1.4610	2.1407	0	0.0000	0.0000	4.0000	5
在线看房	397	1.3098	2.0603	0	0.0000	0.0000	4.0000	5
房屋维修	397	1.4458	2.1142	0	0.0000	0.0000	4.0000	5
旅游专线	397	1.1184	1.9316	0	0.0000	0.0000	3.0000	5
20 秒景点入园	397	1.0756	1.8895	0	0.0000	0.0000	2.5000	5
30 秒酒店入住	397	1.1814	2.0056	0	0.0000	0.0000	3.0000	5
总体满意度	397	4.2443	0.6577	2	4.0000	4.0000	5.0000	5

随机森林模型的验证集或测试集的评估结果见表 2。由表 2 可知, 主要指标 MSE、RMSE 和 MAE 分别表示均方误差、均方根误差和平均绝对误差。上述三个指标的预测值与真实值完全吻合时等于 0, 即完美模型; 预测值和真实值的误差越大, 三个指标的数值越大。从表 2 可以看到三个指标的数值相对较小, 说明预测值和真实值的误差较小, 说明模型的验证集或测试集的评估结果良好, 模型的准确性比较高。

Table 2. Test data evaluation results table**表 2.** 测试数据评估结果表

评价指标	评价结果
MSE	0.2098
RMSE	0.4580
MAE	0.2432
R ²	0.4932
MAPE	6.0784

模型概要特征表见表 3。由表 3 可知，构成训练模型各个特征所占具体权重。权重为负值表示该特征对整体满意度造成负向影响，权重为正值则表示该特征对整体满意度造成正向影响。权重的绝对值越大，表示该特征对整体满意度的影响越大。

Table 3. Model feature profile table**表 3.** 模型特征概要表

特征名称	权重
市民卡	0.0102
生活缴费	0.2215
证件办理	0.0192
健康码	0.0729
舒心就医	0.0694
医保查询	0.1898
停车先离场	0.0216
车位查询	0.0648
交通信息查询	0.1965
公积金	0.0264
在线看房	0.0168
房屋维修	0.0209
旅游专线	0.0197
20 秒景点入园	0.0249
30 秒酒店入住	0.0254

为了直观表现各个特征的影响程度，模型特征概要条形图见图 3。生活缴费、交通信息查询和医保查询这三个模块对用户的总体满意度有明显的正向影响，健康码、舒心就医和车位查询这三个模块对用户的总体满意度有相对明显的正向影响。

该结果意味着上述模块对总体满意度的正向影响较大。因此，从城市大脑建设的角度出发，坚持上述模块的服务质量有利于提高城市大脑的总体满意度。同时，市民卡、在线看房、旅游专线等权重较低的模块可以视为城市大脑今后开发完善的重要空间，为城市大脑未来的发展提供方向上的指导。

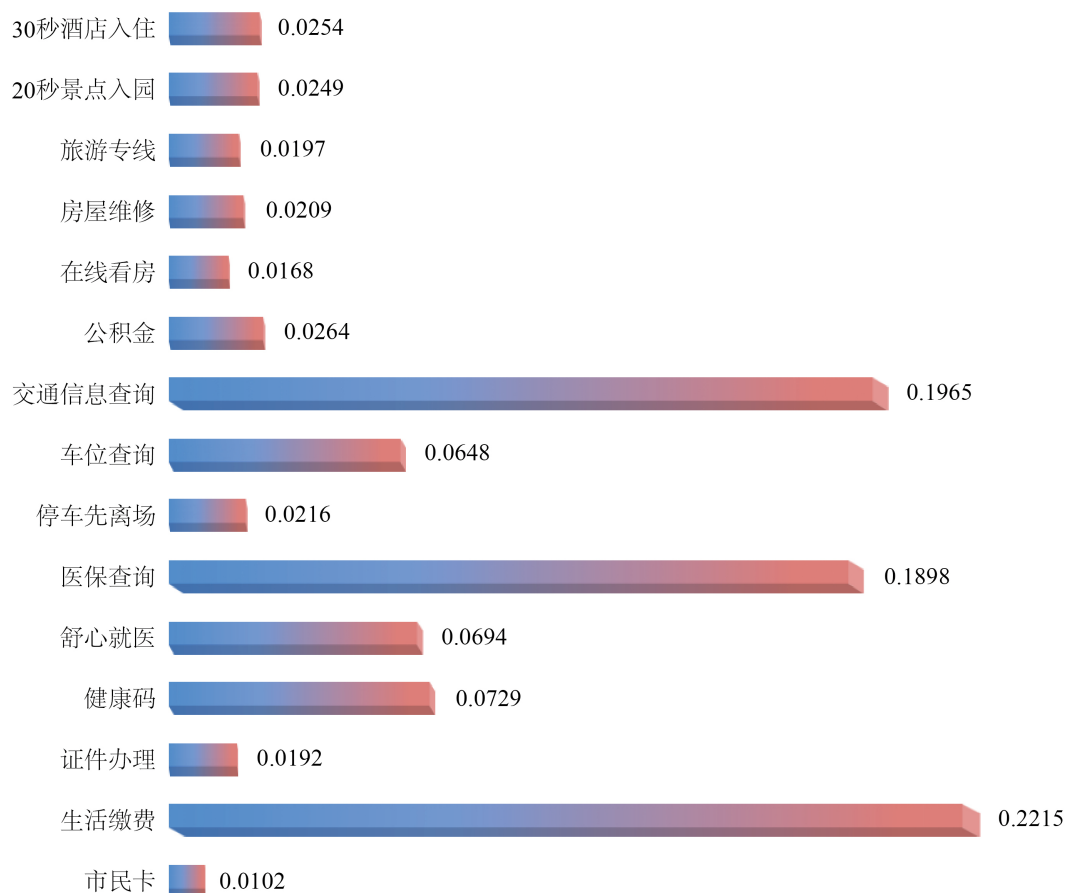


Figure 3. Model feature profile bar chart
图 3. 模型特征概要条形图

4. 结论

在了解现状方面，居民了解城市大脑的途径多样化，其中以网络信息为主，社区的宣传和推广仍存在较大空间。居民对城市大脑有一定了解，但老年群体了解程度偏低，与其他年龄层群体存在明显差异。

在使用现状方面，被调查者使用卫生健康和便民服务模块人数占比较多。特定人群对交通出行、房屋管理和文化旅游模块由对应需求，如 18~44 岁人群作为主要租房购房人群，对房屋管理模块的关注程度更高。使用频率集中在每周 1~3 次，低使用和高使用频率情况较少。

在居民满意度方面，市民卡和证件办理评价最高，而交通信息查询评价最低。交通信息查询模块及时性和稳定性有所欠缺，导致评分受到一定程度影响。总体满意度主要受文化旅游、交通出行和房屋管理因子的影响，便民服务因子和卫生健康因子的作用虽然有一定差距，但也不能忽视。

5. 对策建议

针对上述研究结果，同时结合不同群体特点，本文针对主管部门、基层社区和城市居民分别提出如下对策建议。

1) 主管部门要形成有力技术支撑，推动基础设计建设。

主管部门应将数据安全作为城市大脑建设重点之一。有关技术部门应提高重点保护意识，研发或引进先进数据安全架构技术，为数据安全加上“防护罩”。推动完善智慧化基础设施建设，提高数据处理

水平,有效整理应用数据资源[6]。要加快完善城市大脑智能基础设施建设,保证运行准确率,并将数据进行深度有效运用。要做到城市精细化管理,保证即时、全量、全网、全视频,避免感而不知、盲人摸象、雾里看花和数据资源浪费等情况[7]。

2) 基层社区要加强对内宣传力度,建设完善社区微脑[8]。

社区应主动担起宣传责任,通过宣讲会、宣传栏、上门访谈等方式积极宣传城市大脑。要定期定时向居民普及城市大脑知识,积极帮助社区中部分弱势群体解决问题。将城市大脑宣传通俗化,提高城市大脑普及率与使用率。建设完善社区“微脑”,要依托社区数字化平台和线下社区服务机构,建设智慧服务圈,不断提高服务水平。同时可进一步加快电子政务向社区推进,提高政府办事效率和服务能力。充分体现以人为本、服务民生的宗旨,打造信息畅通、管理有序、服务完善、民生与人际关系和谐的现代化社区[9]。

3) 城市居民要配合城市大脑建设,积极发表反馈建议。

在了解城市大脑的基本知识和相关应用场景后,居民可以借助 APP、小程序等工具使用杭州城市大脑的功能。在方便自己生活的前提下提高使用城市大脑部分模块的能力。居民在使用杭州城市大脑的部分功能时要善于发现其功能不完善的地方或者与现实应用场景不匹配的地方,并向社区或相关主管部门发表反馈与建议,为杭州城市大脑的建设贡献一份力量。

致 谢

感谢各位指导老师在调查方案、调查报告方面给予本项目组的指导。在问卷调查的过程中,435位接受问卷调查的被调查者积极配合本项目组的工作,让本项目组对杭州城市大脑建设现状有更加深入的了解。在此衷心感谢大家的配合,最后也对本项目组成员的辛苦努力表示感谢。

基金项目

本文为浙江省第十届大学生统计调查方案设计大赛参赛作品部分内容;本项目受到2022年浙江省大学生科技创新活动计划暨新苗人才计划项目(2022R407A010)资助。

参考文献

- [1] 杭州:“城市大脑”提升治理效能[J]. 中国建设信息化, 2021(15): 40-41.
- [2] 风笑天. 现代社会调查方法[M]. 第六版. 武汉: 华中科技大学出版社, 2021: 105-157.
- [3] 秒懂杭州城市大脑的48个应用场景[J]. 杭州, 2020(9): 44-49.
- [4] 董师师, 黄哲学. 随机森林理论浅析[J]. 集成技术, 2013, 2(1): 1-7.
- [5] 李欣海. 随机森林模型在分类与回归分析中的应用[J]. 应用昆虫学报, 2013, 50(4): 1190-1197.
- [6] 汪楠, 鲁斐栋, 张心怡. 创新协同运行机制提升城市治理能力——以杭州“城市大脑”为例[J]. 时代建筑, 2021(4): 59-61.
- [7] 王金宝. 浅谈城市大脑与智慧城市发展趋势[J]. 自动化博览, 2020(5): 58-64.
- [8] 李文钊. 数字界面视角下超大城市治理数字化转型原理——以城市大脑为例[J]. 电子政务, 2021(3): 2-16.
- [9] 王延隆, 孙孟瑶. 城市大脑框架下社区小脑建设的思考——以杭州市为例[J]. 创意城市学刊, 2020(3): 157-164.