

DOI:10.11918/j.issn.0367-6234.201805161

一种轻量级梯度提升机的交通模式识别

王 璞,刘 洋,黄智仁

(中南大学 交通运输工程学院,长沙 410000)

摘 要:为研究居民出行轨迹的不同交通方式,基于已知出行交通方式的 GPS 轨迹信息,构建基于轻量级梯度提升机(LightGBM)的分类模型,对居民出行 GPS 轨迹进行交通方式分类. 首先提取轨迹的各种基础特征,并进一步结合公交网络地理信息引入 Fréchet 距离进行特征提取,之后对特征进行归一化处理并基于决策树模型对特征进行筛选,最后对筛选后的特征进行模型训练与预测,并通过五折交叉验证方式获取了稳定的预测结果. 结果表明:公交网络地理特征能够对模型预测精度进行有效提高,提出的 GPS 轨迹交通方式识别方法可达 90%左右的精确度,优于各种机器学习分类模型.

关键词:城市交通;轻量级梯度提升机;GPS 轨迹;特征提取;交通方式识别

中图分类号: U495 **文献标志码:** A **文章编号:** 0367-6234(2019)09-0096-07

Transportation modes recognition using a Light Gradient Boosting Machine

WANG Pu, LIU Yang, HUANG Zhiren

(School of Traffic and Transportation Engineering, Central South University, Changsha 410000, China)

Abstract: To investigate different traffic modes for resident's travel trajectories, a classification model was constructed based on Light Gradient Boosting Machine (LightGBM) to categorize transportation modes according to resident's GPS trajectories. First, basic trajectory features were extracted, and then more features were obtained using geographic information of public transit network (i.e., Fréchet distance). Subsequently, the features were normalized and screened by the decision tree model. Finally, the screened features were trained and predicted by the model, and a stable prediction result was attained with a five-fold cross-validation method. Results show that geographic information of public transit network could optimize the model's prediction accuracy. The proposed GPS trajectory recognition method achieved an accuracy of about 90%, which is superior to other machine learning classification models.

Keywords: urban traffic; LightGBM; GPS trajectories; feature extraction; traffic mode recognition

随着大数据时代的发展,各种手机软件通过采集用户的 GPS 数据能够广泛收集居民的出行信息,通过这些数据不仅能够有效揭露居民的旅行模式^[1-2],了解居民的出行活动^[3-5],也能够为城市的交通规划和管理提供重要的参考^[6]. 然而,通过不同交通方式出行的居民其出行模式存在着显著的差异^[7],例如:小汽车出行通常会覆盖较大的范围并伴随着较快的速度,但是步行则恰好相反. 因此,不同交通出行方式的混淆会对更深一步的挖掘居民的出行模式造成巨大的干扰.

近年来科研学者针对混合多种交通出行模式的 GPS 数据进行了大量的研究并取得了一定的研究成果. 文献[8]利用隐马尔科夫模型对公交车、小汽车、自行车、电动车、火车、地铁、步行、摩托车等 8 种

轨迹进行了识别. 文献[9]基于手机的 GPS 与加速传感器数据,构建了一种多变量参数模型,在对步行和自行车轨迹的识别上能够达到 90%以上的精度. 文献[10]通过结合地理信息数据,提出了一种 GIS 算法,实现了对步行、小汽车、公交车、地铁和通勤铁路 5 种交通方式的轨迹识别. 随着人工智能的发展,各种机器学习和深度学习算法也被广泛应用于交通轨迹的分类,文献[11]基于支持向量机(SVM),文献[12]基于神经网络(NN)均能对轨迹实现较好的分类预测. 文献[13]对支持向量机、随机森林、K-近邻、决策树等机器学习方法进行了对比. 文献[14]使用极端梯度提升模型(XGBoost),并结合全局特征与局部特征对轨迹进行了特征提取,实现了良好的轨迹识别效果. 深度学习诸如卷积神经网络(CNN)^[15]、循环神经网络(RNN)和深度神经网络(DNN)^[16]等均能对 GPS 轨迹进行有效的交通模式分类和识别. 但是现有研究在特征提取时大部分未能考虑地理信息,导致运行状态相似的公交

收稿日期: 2018-05-29

基金项目: 国家自然科学基金(61473320, 71871224)

作者简介: 王 璞(1983—),男,教授,博士生导师

通信作者: 王 璞, wangpu@csu.edu.cn

车与小汽车轨迹较易混淆. 公交车的运行轨迹与公交线路息息相关,本文基于构建的轻量级梯度提升机(LightGBM)模型,引入 Fréchet 距离,结合公交线路与公交站点信息进行特征提取,在提高整体轨迹识别效果的同时,也能够有效改善公交轨迹与小汽车轨迹的识别精度.

1 数据描述与处理

1.1 数据描述

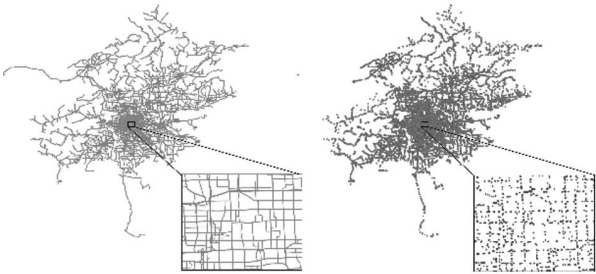
本文研究所涉及的数据包括被标签(交通模式已知)的 GPS 轨迹数据和北京市公交网与公交站点数据. GPS 数据为在 2007 年到 2012 年收集到的 182 位志愿者的 GPS 轨迹数据,其中 73 位志愿者的数据被标签记录,包括步行、自行车、公交车、小汽车、地铁等多种交通出行方式. 一条轨迹由多个 GPS 记录点组成,形式为 $\{r_1, r_2, r_3, \dots, r_n\}$, 其中 r_i 表示每一个 GPS 记录点,每一个 GPS 记录点记录用户其所处位置的经度、纬度、海拔高度以及时间等信息,其记录格式见表 1,其中 91.5% 的数据记录为每 1~5 s 或每 5~10 m 一个 GPS 记录点. 数据获取自微软亚洲研究院官网的开源资源^[17-19]. 北京市公交网与公交站点数据获取自 2013 年,包括 1 543 条公交线路与 42 161 个公交站点,公交线路由线路上的节点连接而成,形式为 $\{b_1, b_2, b_3, \dots, b_m\}$, 公交线路节点与公交站点以经度坐标和纬度坐标的形式记录其地理位置,北京市公交线路和站点分布如图 1 所示^[20].

表 1 GPS 记录点的数据格式

Tab.1 Data format of GPS record point

编号	内容	说明
1	GPS 记录点 ID	GPS 记录点的编号
2	时间	GPS 记录点的时刻记录
3	经度	GPS 记录点位置的经度坐标
4	纬度	GPS 记录点位置的纬度坐标
5	海拔高度	GPS 记录点位置的海拔高度坐标

本文研究范围为北京市(东经 115.4°~117.6°, 北纬 39.4°~41.1°),为了减小随机误差,首先对小于 10 个 GPS 轨迹点的轨迹进行剔除,通过筛选,共获得北京市范围内 7 类共 6 744 条被标签的轨迹(分别为步行 2 007 条,自行车 1 366 条,公交车 1 652 条,小汽车 1 118 条,地铁 563 条,跑步 2 条,火车 36 条),平均每条轨迹包含 495 个 GPS 记录点. 可知跑步以及火车出行的数据量较少,故本研究只针对步行、自行车、公交车、小汽车和地铁 5 类出行轨迹,共 6 707 条.



(a) 公交线路 (b) 公交站点

图 1 北京市公交线网与公交站点分布

Fig.1 Distribution of bus lines and stations in Beijing

1.2 特征提取

1.2.1 基本信息

对于每条被标签的 GPS 轨迹,需要提取轨迹的属性特征进行模型训练,本文研究所涉及的基本特征包括速度、加速度、转角转变量、停留率、速度转变率、转角转变率等,分别定义如下.

速度是各种交通方式最基本的行驶特征,也是区分不同交通模式的最关键指标,一条轨迹中每两个连续 GPS 记录点可计算一次瞬时速度,各类轨迹不同的分位速度具有不同的特征,本文对一条轨迹从 5 分位速度至 95 分位速度之间的每 5 分位速度值进行提取,共获得一条轨迹的 19 个速度特征. 速度 v_i 的计算公式为

$$v_i = \frac{d}{t}.$$

式中: v_i 为用户在轨迹记录点 r_i 处的瞬时速度, d 为 r_i 与 r_{i+1} 间的实际地理距离, t 为 r_i 与 r_{i+1} 间的间隔时间.

加速度是一项重要的行驶特征,每两个连续瞬时速度可以进行一次加速度的计算,同样对轨迹的每 5 分位加速度进行提取,共获得一条轨迹的 19 个加速度特征. 加速度 a_i 的计算公式为

$$a_i = \frac{v_{i+1} - v_i}{t}.$$

式中 a_i 为用户在轨迹记录点 r_i 处的瞬时加速度.

转角转变量也是轨迹的一项重要特征,如图 2(c) 所示,公交车、小汽车和地铁的转角转变量较小,而步行的转角转变量则较大. 每相邻两个 GPS 记录点可以计算一次转角转变量,同样对轨迹的每 5 分位转角转变量进行提取,共获得一条轨迹的 19 个转角转变量特征. 转角转变量 h_i 的计算公式为

$$h_i = |H_{i+1} - H_i|.$$

式中 H_i 为用户在轨迹记录点 r_i 处的行驶方向与正北方向的夹角.

停留率是一条轨迹在单位距离内轨迹点瞬时速度

小于某个特定阈值的轨迹点数量,其分布如图 2(d)所示,小汽车和地铁的停留率较小,步行的停留率较大. 停留率 R_s 的计算公式为

$$R_s = \frac{|P_s|}{D}$$

式中: $|P_s|$ 为轨迹中速度小于某个特定阈值的轨迹点的数量,阈值取为 $3.4 \text{ m/s}^{[18]}$, D 为轨迹的路径长度.

速度转变率是一条轨迹在单位距离内速率变化大于某个特定阈值的轨迹点数量,其分布如图 2(e)所示,小汽车和地铁的速度转变率较小,步行的速度转变率较大. 速度转变率 R_{vc} 的计算公式为

$$R_{vc} = \frac{|P_v|}{D}$$

式中: $|P_v|$ 为轨迹中速率变化大于某个特定阈值的轨迹点的数量,阈值取为 $0.26 \text{ m/s}^{2[18]}$, 速率变化

的公式为

$$p_i = \frac{|v_i - v_{i+1}|}{v_i}$$

转角转变率是一条轨迹在单位距离内转角转变量大于某个特定阈值的轨迹点数量,其分布如图 2(f)所示,同样小汽车和地铁的转角转变率较小,步行的转角转变率较大. 转角转变率 R_{hc} 的计算公式为

$$R_{hc} = \frac{|P_c|}{D}$$

式中 $|P_c|$ 为轨迹中转角转变量大于某个特定阈值的轨迹点的数量,阈值取为 $19^\circ^{[18]}$.

其中 50 分位速度、50 分位加速度、50 分位转角转变量、停留率、速度转变率与转角转变率的轨迹特征分布如图 2(a)~2(f)所示.

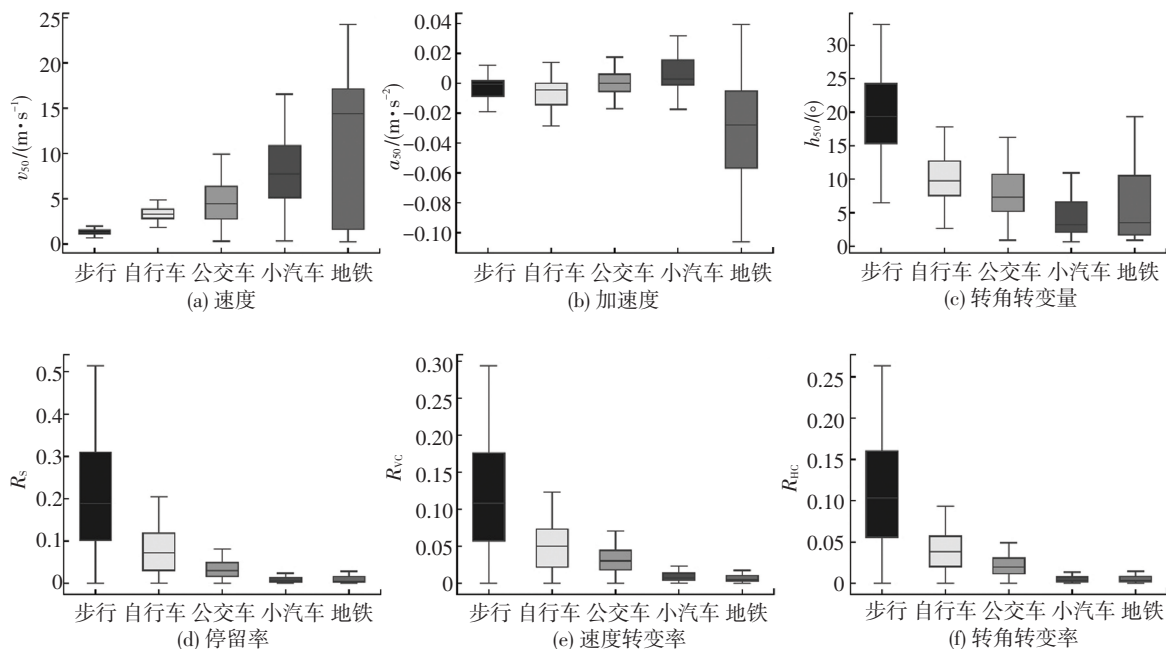


图 2 轨迹基本特征分布

Fig.2 Basic features distribution

1.2.2 公交地理信息

为了对公交轨迹实现更好的识别,融入公交线路与公交站点信息以获得与公交地理信息相关的特征,本文共获得 3 个公交地理特征,轨迹起点距公交站点距离、轨迹终点距公交站点距离和轨迹同公交线路的 Fréchet 距离,分别定义如下.

轨迹起点距公交站点距离 d_{if} : 轨迹的第一个 GPS 记录点与最近公交站点的地理距离,其分布如图 3(a)所示,可以看出公交轨迹的距离相对较小.

轨迹终点距公交站点距离 d_{te} : 轨迹的最后一个 GPS 记录点与最近公交站点的地理距离,其分布如

图 3(b)所示,同样公交轨迹的距离相对较小.

轨迹与公交线路的 Fréchet 距离 $d_{\text{Fréchet}}(P, Q)$: Fréchet 距离是一种空间路径相似性描述的方法,其重点考虑曲线间的空间距离,对有着空间时序的曲线通过 Fréchet 距离能够很好地描述曲线间的空间相似程度^[21]. 对于两条由离散的点连接而成的曲线 $P\{r_1, r_2, r_3, \dots, r_n\}$ 和 $Q\{b_1, b_2, b_3, \dots, b_m\}$, Fréchet 距离的计算方式为

$$d_{\text{Fréchet}}(P, Q) = \inf_{i \in [1, n]} \max_{j \in [1, m]} \{d(r_i, b_j)\}.$$

式中: P 为用户轨迹, Q 为公交线路, $r_i \in \{r_1, r_2, r_3, \dots, r_n\}$, $b_j \in \{b_1, b_2, b_3, \dots, b_m\}$, b_j 为公交线

路的节点集合, $d(r_i, b_j)$ 为两 GPS 点的距离度量函数。

对于 Fréchet 距离的计算通过如下步骤: 1) 对轨迹与公交线路进行匹配, 分别筛选距轨迹起点与终点 1 500 m 之内(图 4 中圆形区域)的公交线路集合 B_f 与 B_e , 获取 B_f 与 B_e 中共有的公交线路 $B_f \cap B_e$ 作为候选公交线(图 4 中的公交线路 a, b, c)。2) 对于每条候选公交线, 截取距轨迹起点最近的公交线节点和距轨迹终点最近的公交线节点的公交线范围作

为候选截断公交线。3) 分别计算轨迹与候选截断公交线路的 Fréchet 距离, 选定 Fréchet 距离最小值作为该轨迹的 Fréchet 距离特征(图 4 中的公交线路 a 为 Fréchet 距离最小值所对应的公交线)。

不同轨迹的 Fréchet 距离分布如图 3(c) 所示, 可以看出步行、自行车与公交车轨迹的 Fréchet 距离明显较小, 而小汽车与地铁轨迹的 Fréchet 距离则相对较大。

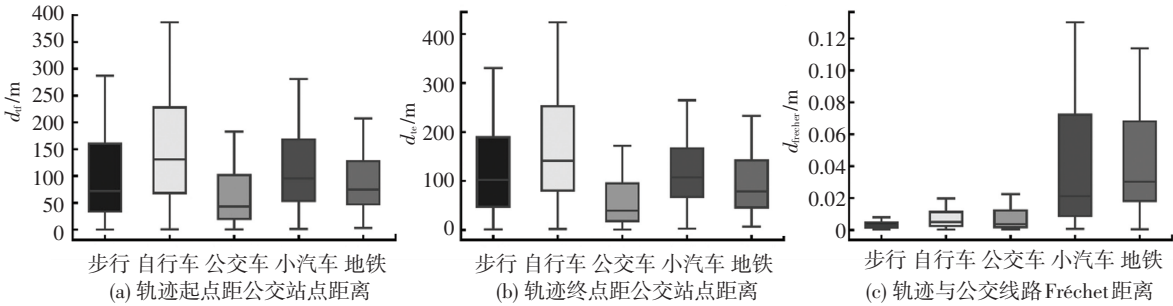


图 3 轨迹地理特征分布

Fig.3 Geographic features distribution

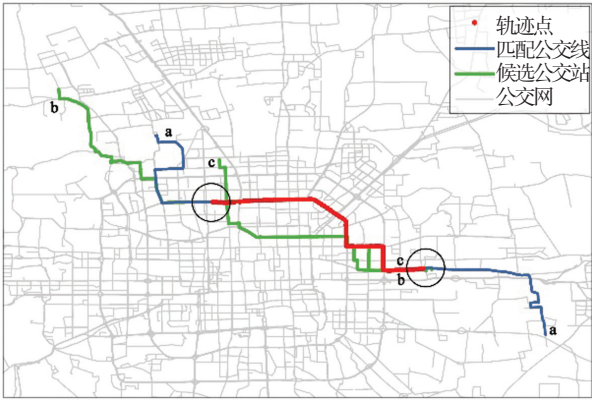


图 4 轨迹与公交线路 Fréchet 距离计算示意

Fig.4 Calculation of Fréchet distance between trajectory and bus line

2 GPS 轨迹交通模式识别模型

2.1 模型介绍

轻量级梯度提升机 (LightGBM) 基于梯度提升决策树 (GBDT), 是一种基于直方图 (Histogram) 的决策树算法, 其通过加法模型和前向分步算法实现学习的优化, 较比传统的 GBDT, LightGBM 通过直方图方式对连续特征值进行分段, 能够实现更快的模型训练速度并节省内存, 并且通过 leaf-wise 的生长策略对树进行生长, 并限制树的深度以防止过拟合, 能够有效提高模型预测的准确率。LightGBM 是微软开发的一种开源分类预测算法, 本研究通过调用 python 的 lightgbm 模块实现。LightGBM 的具体实现流程如下。

模型训练阶段:

输入: 训练数据 (X_{train}, Y_{train})

1、确定目标函数:

$$obj^t = \sum_{i=1}^n l \cdot (y_i, \hat{y}_i^{(t)}) + \sum_{i=1}^l \Omega f(i)$$

2、初始化常数: $\hat{y}_i^{(0)} = 0$

3、While 未达到迭代次数或预设的终止条件:

4、计算目标函数, 确定梯度下降方向与步长, 更新决策树 $f(t)$

4.1、对特征值构建直方图, 计算各个切分点的增益, 寻找最优分割点

4.2、带有深度限制的 Leaf-wise 树生长策略, 限制树的深度与叶节点的最小样本数

5、更新训练结果:

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i)$$

模型测试阶段:

输入: 测试数据 X_{test}

6、加权所有决策树的结果: $\hat{y}^t = \sum_{k=1}^K f_k(x_i)$

输出: 测试结果 Y_{test}

其中: x_i 为某个训练样本或测试样本的特征值, y_i 为训练样本或测试样本的目标值, n 为训练样本的个数, $\hat{y}_i^{(t)}$ 为第 t 轮迭代得到的 y_i 的预测值, $l \cdot (y_i, \hat{y}_i^{(t)})$ 为损失函数, $f(i)$ 为第 i 轮迭代的决策树, $\sum_{i=1}^l \Omega f(i)$ 为正则化项, K 为最终迭代次数。

2.2 训练方式与结果评估

首先对轨迹特征进行标准化处理,将所有特征值缩放到 $[0,1]$ 之间,以避免各特征值量级的不同而导致的误差,公式为

$$X_{\text{std}} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}}$$

式中: X 为所有轨迹的某一特征值集合, X_i 为轨迹 i 的原特征值, X_{std} 为轨迹 i 在进行特性标准化处理后的特征值.

然后通过决策树模型计算轨迹特征(每条轨迹 63 个特征)与轨迹标签(交通模式)的相关程度,并对每个特征的重要性进行排序(相关程度从大到小排序),筛选重要性排名前 n 的特征作为模型的输入,用于模型训练与预测. 对于模型学习阶段的随机误差,采用 5 折交叉验证的方式分别构建训练集与测试集并重复运算 100 遍以保证模型学习的稳定性.

对预测结果精确程度的度量,考虑 4 个指标,准确率(A_{CC}),精确率(P),召回率(R),与调和平均值($F1$),其中准确率为正确分类样本所占的比例,精确率为预测结果为正样本中正确分类样本所占的比例(即查准率),召回率为实际为正样本中被正确分类样本所占的比例(即查全率),调和平均值为综合考虑对精确率与召回率的衡量. 计算方式分别为

$$A_{\text{CC}} = \frac{T_p + T_N}{T_p + T_N + F_p + F_N},$$
$$P = \frac{T_p}{T_p + F_p},$$
$$R = \frac{T_p}{T_p + F_N},$$

模式	样本数					$A_{\text{CC}} / \%$	$R / \%$	$F1 / \%$	$P / \%$
	步行	自行车	公交车	小汽车	地铁				
步行	374.4	8.7	12.3	1.3	4.7	91.50	93.28	92.38	95.40
自行车	12.4	252.0	7.8	0.8	0.2	94.48	92.25	93.35	94.29
公交车	11.4	4.5	286.0	25.5	3.0	84.26	86.55	85.39	91.24
小汽车	3.1	1.1	30.1	186.1	3.2	85.17	83.22	84.18	93.33
地铁	7.9	0.5	3.2	4.8	96.2	89.61	85.48	87.50	97.95

由表 2 可知步行与自行车出行两种轨迹较易区分,预测精度(精确率和召回率)可达 90% 以上,但是公交车和小汽车两种交通轨迹较易混淆,预测精度相对稍低. 公交车一般沿着公交线路行驶,加入能够突出公交轨迹的公交地理特征后,所得预测结果见表 3.

$$F1 = \frac{2 T_p}{2 T_p + F_p + F_N}.$$

式中: T_p 为将正类预测为正类的样本数量, F_N 为将正类预测为负类的样本数量, F_p 为将负类预测为正类的样本数量, T_N 为将负类预测为负类的样本数量.

对经由决策树模型筛选的不同特征数量对其预测准确率进行统计,发现当选择特征数量为 55 个时,预测效果达到较好,如图 5 所示,故最终选定模型的特征数 $n = 55$.

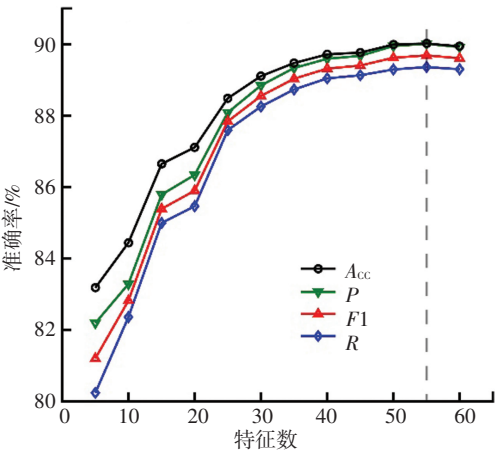


图 5 特征数量与预测精度的关系分布

Fig.5 Relationship between characteristic quantity and prediction accuracy

2.3 结果分析

2.3.1 LightGBM 结果

对于未考虑地理特征(d_{lf} 、 d_{le} 与 $d_{\text{Fréchet}}(P,Q)$), 所得预测结果见表 2.

表 2 未考虑地理特征的预测结果统计

Tab.2 Prediction results without reference to geographical features

表 3 考虑地理特征的预测结果统计

Tab.3 Prediction results with reference to geographical features

模式	样本数					$A_{cc} / \%$	$R / \%$	$F1 / \%$	$P / \%$
	步行	自行车	公交车	小汽车	地铁				
步行	375.7	8.5	12.4	1.8	3.0	92.17	93.60	92.88	95.70
自行车	12.2	252.2	7.6	0.8	0.3	94.63	92.31	93.46	94.62
公交车	11.0	4.0	289.0	23.1	3.3	86.04	87.48	86.76	92.19
小汽车	3.2	1.1	24.5	191.7	3.2	86.10	85.73	85.92	94.18
地铁	5.6	0.6	2.4	5.3	98.7	90.98	87.67	89.30	98.24

表 4 不同模型结果对比

Tab.4 Prediction results of different models

模型	$A_{cc} / \%$	$P / \%$	$R / \%$	$F1 / \%$
LR	76.36	74.33	70.04	71.28
KNN	85.24	85.14	82.73	83.73
SVM	76.87	77.29	68.40	69.81
RF	88.53	88.46	87.54	87.95
GBDT	88.88	88.65	88.00	88.29
MLP	84.13	83.00	81.01	81.80
LightGBM	90.02	90.01	89.36	89.69

可以看出,较比其他机器学习模型,LightGBM 模型能够实现更为有效的预测,其预测精度可达 90%左右,其他模型如 GBDT 模型和 RF 模型相对次之,之后是 KNN 模型和 MLP 模型,LR 模型和 SVM 模型则预测精度相对较差。

3 结 论

1) 基于被交通模式标记的 GPS 轨迹,利用轻量级梯度提升机(LightGBM)分类模型对交通出行轨迹进行了识别,考虑速度、加速度、转角等多种因素,并结合公交地理信息,引入 Fréchet 距离对 GPS 轨迹进行特征提取,实现了对 GPS 轨迹的交通模式分类。

2) 加入公交地理信息能够对公交轨迹与小汽车轨迹实现更有效地识别,公交轨迹和小汽车轨迹的预测精度均能提高 2%左右。

3) 同等条件下较比其他机器学习分类模型,LightGBM 能够实现更为有效的轨迹分类预测,整体预测精度达到 90%(整体预测精度即综合考虑准确率、精确率、召回率与调和平均值 4 种指标的预测精度,其中准确率为 90.02%,精确率为 90.01%,召回率为 89.36%,调和平均值为 89.69%),其中步行和自行车出行两类轨迹预测精度可达 90%以上。

4) 本研究能够对由手机软件收集的用户 GPS 轨迹实现有效的交通模式分类,从而更为有效地了解用户的出行行为与出行特性,能够为城市交通的

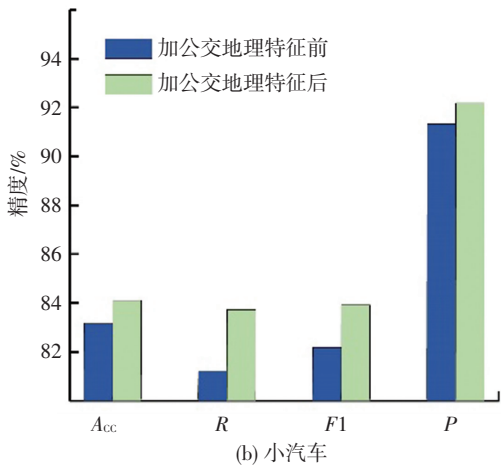
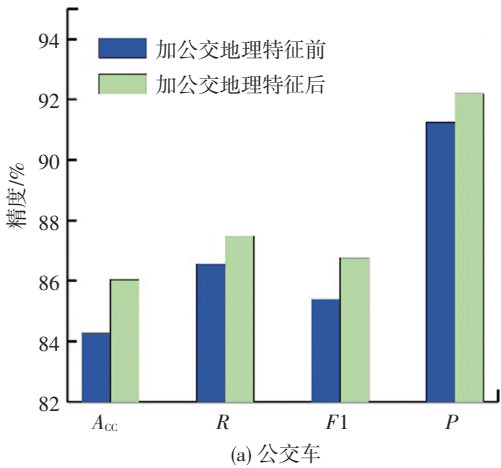


图 6 公交地理特征加入前后精度对比
Fig.6 Comparison of prediction accuracy with and without reference to geographical features

2.3.2 不同模型结果对比

将 LightGBM 模型与其他机器学习模型:逻辑回归(LR)、K-近邻(KNN)、支持向量机(SVM)、随机森林(RF)、梯度提升决策树(GBDT)、多层感知神经网络(MLP)的预测效果进行对比,各种模型的特征选择与训练方式均保持一致,采用 5 折交叉验证,重复 100 次后取平均结果见表 4。

规划及商业布局提供良好的基础信息. 与公交轨迹相似, 地铁轨迹是沿着地铁线运行, 将来可结合地铁线路信息进行更深入的特征提取以对该轨迹分类方法实现进一步的优化.

参考文献

- [1] SADEGHVAZIRI E, ROJAS M B, JIN X. Exploring the potential of mobile phone data in travel pattern analysis[J]. Transportation research record, 2016, 2594(1): 27
- [2] 王璞, 黄智仁, 龚航. 大数据时代的交通工程[J]. 电子科技大学学报, 2013, 42(6): 806
WANG Pu, HUANG Zhiren, GONG Hang. Transportation engineering in the big data era[J]. Journal of University of Electronic Science and Technology of China, 2013, 42(6): 806
- [3] WANG P, HUNTER T, BAYEN A M, et al. Understanding road usage patterns in urban areas[J]. Scientific Reports, 2012, 2: 1001
- [4] HUANG Z, WANG P, ZHANG F, et al. A mobility network approach to identify and anticipate large crowd gatherings[J]. Transportation Research Part B: Methodological, 2018, 114: 147
- [5] ZHENG Z, HUANG Z, ZHANG F, et al. Understanding coupling dynamics of public transportation networks[J]. EPJ Data Science, 2018, 7(1): 23
- [6] HE K, XU Z, WANG P, et al. Congestion avoidance routing based on large-scale social signals[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 17(9): 2613
- [7] 胡晓伟, 王健, 孙广林. 有限理性下出行者方式选择行为[J]. 哈尔滨工业大学学报, 2011, 43(12): 114
HU Xiaowei, WANG Jian, SUN Guanglin. Traveler's mode choice behavior analysis under bounded rational[J]. Journal of Harbin Institute of Technology, 2011, 43(12): 114
- [8] WIDHALM P, NITSCHKE P, BRANDIE N. Transport mode detection with realistic smartphone sensor data[C]//International Conference on Pattern Recognition IEEE. Tsukuba: IEEE Press, 2012
- [9] NITSCHKE P, WIDHALM P, BREUSS S, et al. A strategy on how to utilize smartphones for automatically reconstructing trips in travel surveys[J]. Procedia-Social and Behavioral Sciences, 2012, 48: 1033
- [10] GONG H, CHEN C, BIALOSTOZKY E, et al. A GPS/GIS method for travel mode detection in New York City[J]. Computers, Environment and Urban Systems, 2012, 36(2): 131
- [11] BOLBOL A, CHENG T, TSAPAKIS I, et al. Inferring hybrid transportation modes from sparse GPS data using a moving window SVM classification[J]. Computers, Environment and Urban Systems, 2012, 36(6): 526
- [12] GONZALEZ P A, WEINSTEIN J S, BARBEAU S J, et al. Automating mode detection for travel behaviour analysis by using global positioning systems-enabled mobile phones and neural networks[J]. IET Intelligent Transport Systems, 2010, 4(1): 37
- [13] JAHANGIRI A, RAKHA H A. Applying machine learning techniques to transportation mode recognition using mobile phone sensor data[J]. IEEE Transactions on Intelligent Transportation Systems, 2015, 16(5): 2406
- [14] XIAO Z, WANG Y, FU K, et al. Identifying different transportation modes from trajectory data using tree-based ensemble classifiers[J]. ISPRS International Journal of Geo-Information, 2017, 6(2): 57
- [15] DABIRI S, HEASLIP K. Inferring transportation modes from GPS trajectories using a convolutional neural network[J]. Transportation Research Part C: Emerging Technologies, 2018, 86: 360
- [16] ENDO Y, TODA H, NISHIDA K, et al. Classifying spatial trajectories using representation learning[J]. International Journal of Data Science and Analytics, 2016, 2(3/4): 107
- [17] ZHENG Y, ZHANG L, XIE X, et al. Mining interesting locations and travel sequences from GPS trajectories[C]//Proceedings of the 18th International Conference on World Wide Web. Madrid: ACM Press, 2009: 791
- [18] ZHENG Y, LI Q, CHEN Y, et al. Understanding mobility based on GPS data[C]//Proceedings of the 10th International Conference on Ubiquitous Computing. New York: ACM Press, 2008: 312
- [19] ZHENG Y, XIE X, MA W Y. Geolife: a collaborative social networking service among user, location and trajectory[J]. IEEE Data Eng Bull, 2010, 33(2): 32
- [20] LONG Y, THILL J C. Combining smart card data and household travel survey to analyze jobs-housing relationships in Beijing[J]. Computers, Environment and Urban Systems, 2015, 53: 19
- [21] FRECHET M M. Sur quelques points du calcul fonctionnel[J]. Rendiconti del Circolo Matematico di Palermo (1884 - 1940), 1906, 22(1): 1

(编辑 魏希柱)