

文章编号: 2095-2163(2024)02-0162-04

中图分类号: TP183

文献标志码: A

基于元学习方法的道路车速预测模型

李子承, 仲 铭

(南京审计大学 商学院, 南京 210029)

摘要: 准确预测交通道路上的车速,有助于提升交通安全、缓解交通拥挤、促进自动驾驶技术的发展。针对这一需求,本文首次将元学习方法应用到道路车速预测这一领域,研究了基于元学习方法的道路车速预测模型。构建滑动时间窗口,通过6个基础预测器进行初步预测,并用元学习方法给这些预测结果分配一组权值,最终形成组合预测。在公开的数据集上进行了实验,结果表明:该元学习器的预测性能优于目前较为先进的基准模型,为道路车速预测模型提供了更好的解决方案。

关键词: 元学习; 车速预测; 城市交通

Road speed prediction model based on Meta learning

LI Zicheng, ZHONG Ming

(School of Business, Nanjing Audit University, Nanjing 210029, China)

Abstract: Accurately predicting vehicle speeds on traffic roads contributes to enhancing traffic safety, alleviating congestion, and promoting the development of autonomous driving technology. In response to this need, this paper, for the first time, applies Meta-learning methods to the field of road speed prediction and investigates a road speed prediction model based on Meta-learning methods. It constructs a sliding time window, conducts preliminary predictions using six basic predictors, and assigns a set of weights to these prediction results using Meta-learning methods to form a combination prediction. Experiments conducted on public datasets show that the predictive performance of this Meta-learner exceeds that of currently advanced benchmark models, offering a superior solution for road speed prediction models.

Key words: Meta-learning; vehicle speed prediction; urban traffic

0 引言

随着现代城市化的发展,城市交通拥堵问题日益严重,交通事故愈发严重。根据2018年《全球道路安全现状》,全球每年因交通事故受伤者人数超过2000万人。面对严峻的道路交通问题现状,如果能对每条道路上的车速进行精准预测,不仅对道路安全起到很好的作用,对交通拥挤与自动驾驶技术的发展也起到支持作用。

近几十年来,人们在车速预测模型方面的研究付出了大量的努力。其模型主要可以分为两大类:一种是基于车路协同模型法,主要是以微观仿真模型来描述交通数据的变化;另一种是基于数据驱动法^[1],其中包括随机预测、浅层学习、深度学习等。由于前者构建模型成本昂贵,且车辆作为人们日常出行的交通工具,通常有较为固定的路线和时间规律,随着时间流逝反映在车速数据上具有规律性的

特征^[2]。在这种情形下,以历史数据建立预测模型,是更为适合的方法。

传统基于数据驱动的车速预测,大多采用统计分析和机器学习方法。早期研究时,参数化方法如自回归移动平均线(ARMA)、卡尔曼滤波器、自回归集成移动平均线(ARIMA)常常被使用。然而,在不稳定的交通条件和复杂的道路中,与非参数化相比,这些方法通常预测性能较差^[3-4]。神经网络(NN)、K-最近邻(KNN)、贝叶斯网络(BN)和支持向量机是非参数化方法的代表,其没有关于基础模型公式的假设和估计模型参数的不确定性,预测精度相对更好。机器学习方法因为优秀的学习能力与适应性,所以逐渐在交通预测领域得到了广泛应用。神经网络通常用于交通流预测^[5],递归神经网络(RNN)因其短期记忆被广泛用于非线性时间序列数据的建模。Tang等^[6]引入了改进的模糊神经网络(FNN),以提高交通流预测的准确性。

作者简介: 李子承(1998-),男,硕士研究生,主要研究方向:时间序列分析及应用; 仲 铭(1999-),男,硕士研究生,主要研究方向:风电预测。

收稿日期: 2023-02-24

哈尔滨工业大学主办 ◆ 科技创见与应用

混合模型结合了不同方法的优点, 进一步提高了预测精度。例如, Ling 等^[7]提出了一种结合马特拉小波和 BP 神经网络的短期交通预测模型; 宋旭东等^[8]提出了基于组合模型的短时交通流预测; 肖宇等^[9]提出了基于 XGBoost 算法的短时出租车预测模型, 并使用基于贝叶斯优化的方法对超参数方法进行优化。

根据 Wolpert 等^[10]提出的“没有免费午餐”定理: “不能保证任何方法, 无论其多么复杂, 对不同的数据集, 总比另一种预测方法表现得更好”。基于此观点, 在混合模型的基础上, 本文进一步提出了基于元学习方法^[11]的道路车速预测模型, 模型使用 6 个不同的基础预测器, 分别对历史数据集进行预测, 元学习器的目标则是确定最佳权重, 组合基础预测器的预测而产生组合预测, 以获得最小的预测误差。

差。

1 方法论

城市交通道路车速预测的元学习方法主要流程为: 首先定义一组基预测器, 以各个基预测器在不同的交通道路车速时间序列数据样本中的表现作为元数据; 然后用深度学习神经网络对元数据进行学习, 得到将道路车速时间序列数据的特征与基预测模型的表现联系起来的元知识; 最后用所学习到的元知识来组合基预测模型进行组合预测。

1.1 基础预测器

根据交通车速时间序列预测问题的特点, 将使用 6 个不同的基预测器。根据各基预测器的特点对其进行组合, 有助于提高元学习组合预测的精度。基预测器及其各自主要优势见表 1。

表 1 六个基预测器的不同优势

Table 1 The strengths of six difference base predictors

模型	优势
整合移动平均自回归模型 (Arima)	模型简单
支持向量机模型 (SVM)	鲁棒性较高, 优秀的泛化能力
岭回归模型	综合权衡方差与误差
随机森林回归模型	泛化能力强, 能够平衡误差
Xgboost 回归模型	正则化项防止过拟合
LightGBM 回归模型	训练时间快, 并行学习, 具有处理大数据的能力

1.2 元学习器

近年来, 特征学习已成为机器学习的一个重要领域, 最成功的特征学习框架采用深度神经网络。其中, 卷积神经网络 (CNN) 能够自动挖掘和产生输入图像或时间序列的深层特征, 并对数据平移、缩放和旋转具有很强的鲁棒性, 在计算机视觉、模式识别、机器学习等问题中表现突出。

本文采用 Ma 等^[12]提出的基于双通道卷积神经网络 (DCCNN) 的元学习器, 两个通道分别以车速时间序列 $[y_{i,t}]_{t=1:T}$ 和车速影响因素 $[X_u]_{t=1:T+H}$ 时间序列构成的多元时间序列为输入。更正式表示为: 对于 d 维的多元时间序列输入 (或前一层的输出) $z \in R^{d \times T}$, 以及一维滤波器 $v \in R^s$, 对于步长 s , 输入 z 的时间步长 t 上的卷积运算 C 定义为式 (1):

$$C(t) = (v * z) = \sum_{j=1}^d \sum_{i=0}^{s-1} v_i z_{j,t-i} \quad (1)$$

两个卷积信道都包含 3 个堆叠的时间卷积块, 用作特征提取。每个卷积块包含卷积层和 ReLU 激

活层 ($\text{ReLU}(u) = \max(0, u)$), 前两个卷积块以挤压和激发层结束。挤压操作探索各个过滤器聚焦特征外的上下文信息, 通过使用全局平均池生成学习特征映射的摘要统计信息。具体而言, 卷积层输出 U 通过时间维度 T 收缩, 以计算汇总统计 \bar{U} 。 \bar{U} 的第 k 个元素计算公式为式 (2):

$$\bar{u}_k = \frac{1}{T} \sum_{t=1}^T u_{k,t} \quad (2)$$

通过挤压操作获取摘要信息后进行激发操作, 其目的是捕获所学习特征之间的依赖关系。为了实现这一点, 采用简单的门机制和 sigmoid 激活函数, 具体操作见式 (3):

$$q = \sigma(w_2 \text{ReLU}(w_1 \bar{U})) \quad (3)$$

式中: σ 是 sigmoid 激活函数, $w_1 \in R^{r \times K}$ 与 $w_2 \in R^{K \times K}$ 是可学习的权重, w_1 为降维层参数, w_2 为增维层参数, r 是门机制的降维率, 可以根据当前输入动态控制信息流。最后, 模块的输出按式 (4) 重新压缩:

$$\tilde{u}_k = q_k \cdot u_k \quad (4)$$

最后一个时间卷积块后是一个全局平均池化层,用于减少网络中的参数数量。利用 softmax 激活的 dense 层权值,对 M 个预测进行加权求和,生成一组组合预测,计算公式为式(5):

$$\hat{y}_{i,T+h} = \sum_m w_i^{(m)} \hat{y}_{i,T+h}^{(m)} \quad (5)$$

其中, $\hat{y}_{i,T+h}^{(m)}$ ($h=1:H, m=1:M$) 是时间序列 i 在时间步 h 上从第 m 个基预测器产生的预测值, w 是从 softmax 层输出的组合权重,满足 $\sum_m w_i^{(m)} = 1$ 。

组合预测使用比例均方误差 (SMSE) 损失函数进行评估,其定义为式(6):

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \frac{\frac{1}{H} \sum_{h=1}^H (\hat{y}_{i,T+h}(\theta) - y_{i,T+h})^2}{S_i} \quad (6)$$

其中, θ 表示网络中的所有参数, S_i 是 M 个基预测器的均方误差值,其具体定义为式(7):

$$S_i = \frac{1}{MH} \sum_{m=1}^M \sum_{h=1}^H (\hat{y}_{i,T+h}^{(m)} - y_{i,T+h})^2 \quad (7)$$

2 实验结果分析

2.1 数据集来源

本文实验评估预测性能选取了洛杉矶公路这一份知名的公开数据集,该数据集是在高速公路上收集的通过环路探测器实时监测洛杉矶各个道路的情况。时间跨度为7天,共有207个探测器,根据固定频率汇总交通速度,时间步长为30分钟。

本文使用一个112小时宽的固定滚动窗口进行估计和预测。每4小时向前移动一次窗口,在168小时的数据样本中生成15个数据槽。前10个数据槽用作元学习的训练数据,后5个数据槽用于测试元学习器的预测性能。对于每个时段的数据,前108小时的数据用于训练基础预测器,后4小时的数据用于评估基础预测器的表现;使用“插值方式”对缺失值进行预处理。

2.2 预测基准

除了使用6个基预测器作为预测基准模型,本文还设计了3种广泛使用的组合预测方法作为基准,以更好地评估元学习方法的预测表现。

(1) 等权重组合 (E1): 使用算术平均值对所有基预测器的预测进行简单平均。

(2) 加权线性组合 (E2): 权重根据基预测器在训练数据上的表现,通过 softmax 函数进行加权。

(3) 基预测器选择后的等效权重组合 (E3): 根

据其在训练数据上的表现,选择最优的4个基预测器,采用等价权重法对所选基预测器的预测进行组合。

2.3 评价指标

恰当的评价指标有利于准确直观地反映模型的预测效果。本文使用平均绝对误差 (MAE) 与对称平均绝对百分比误差 (sMAPE) 两种误差度量来衡量模型的预测性能,其计算公式为:

$$MAE = \frac{1}{NH} \sum_{i=1}^N \sum_{h=T+1}^{T+H} |y_{ih} - \hat{y}_{ih}| \quad (8)$$

$$sMAPE = \frac{1}{NH} \sum_{i=1}^N \sum_{h=T+1}^{T+H} \left| \frac{y_{ih} - \hat{y}_{ih}}{y_{ih} + \hat{y}_{ih}} \right| \quad (9)$$

其中, \hat{y}_{ih} 是预测期 h 中第 i 个预测值; y_{ih} 是对应的真实值; N 是样本中的数量。

2.4 结果分析

洛杉矶公路数据集的基预测器在测试期间的预测表现见表2。表中模型后面的数字表示使用道路速度时间序列滞后的阶数。从表2中可以看出,若为每条道路序列进行单独建模,即独立预测每一条道路上车速的基预测器 (Arima、SVM、Ridge), 效果不如使用所有道路序列一起训练的全局预测模型。显然,在道路车速预测方面,使用全局预测模型会取得更好的预测精度;此外,在车速预测方面,6个基预测器中, Lgb 回归模型表现出了最高的预测精度,而 Arima 表现最差。

表3与图1展示了数据集的组合预测 (E1, E2, E3) 以及元学习器 (Meta) 在测试期间的预测表现,并将基预测器中表现最好的 Lgb 回归模型当作基模型进行对比。结果表明,元学习器的预测精度明显优于基预测器以及3个组合预测。

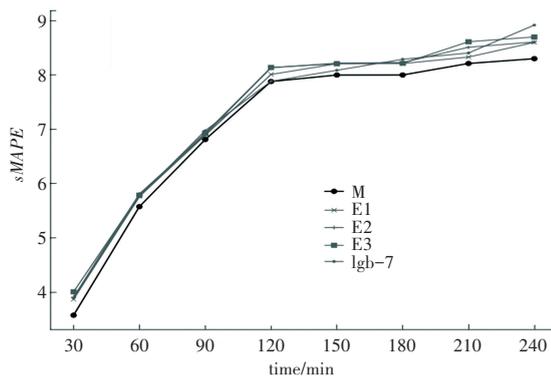


图1 元学习器与组合预测 sMAPE 对比

Fig. 1 Comparison of sMAPE between Meta-learner and combined prediction

表 2 基础预测器在测试集的预测表现

Table 2 Prediction performance of the base predictor on the test set

模型	30 min		60 min		240 min		平均	
	sMAPE	MAE	sMAPE	MAE	sMAPE	MAE	sMAPE	MAE
Arima	6.230	5.981	10.503	10.411	12.051	10.060	9.755	9.116
SVM-1	3.971	3.383	8.072	7.042	8.773	8.272	7.504	6.710
SVM-3	4.000	3.364	8.100	7.084	8.638	8.117	7.496	6.688
Ridge1	4.320	3.599	8.337	7.358	9.429	8.980	7.934	7.128
Ridge3	4.069	3.386	8.185	7.190	9.541	9.115	7.832	7.031
Rf-3	4.335	3.661	8.386	7.473	8.387	7.796	7.683	6.892
Rf-7	4.224	3.580	8.207	8.207	8.315	8.317	7.587	7.587
Xgb-3	4.276	3.575	8.169	7.233	8.074	7.412	7.448	6.627
Xgb-7	4.071	3.377	7.899	6.954	8.114	7.485	7.321	6.494
Lgb-3	4.093	3.396	8.215	7.247	8.042	7.398	7.421	6.581
Lgb-7	3.895	3.261	7.885	6.925	8.087	7.450	7.277	6.458

表 3 三个组合基准模型和元学习器的预测性能

Table 3 Prediction performance of three combined benchmark models and Meta-learner

模型	30 min		60 min		240 min		平均	
	sMAPE	MAE	sMAPE	MAE	sMAPE	MAE	sMAPE	MAE
Meta	3.579	3.115	7.881	6.810	8.001	7.211	6.771	6.215
E1	3.872	3.219	8.218	6.819	8.367	7.225	7.991	6.341
E2	3.913	3.234	8.201	6.833	8.356	7.313	7.879	6.399
E3	4.011	3.301	8.209	6.925	8.441	7.299	7.983	6.454
Lgb-7	3.895	3.261	7.885	6.925	8.087	7.450	7.277	6.458

3 结束语

本文利用一个真实世界交通道路车速的历史数据,首次对元学习方法在车速预测方面进行了应用。该方法使用一种基于卷积神经网络的元学习器,采用有监督学习方法从道路车速时间序列及其影响因素中自动提取特征,并使用混合的基预测器。通过与多个时间序列预测方法在车速预测方面的对比实验表明,元学习方法在预测道路车速方面表现出了更高的性能。

参考文献

[1] 贺正冰,徐瑞康,谢东繁,等. 数据驱动跟驰模型综述[J]. 交通运输系统工程与信息,2021,21(5):102-113.
 [2] ZHANG L, LIU W, QI B. Combined prediction for vehicle speed with fixed route[J]. Chinese Journal of Mechanical Engineering, 2020,33(1):60.
 [3] LI L, HE S, ZHANG J, et al. Short-term highway traffic flow prediction based on a hybrid strategy considering temporal-spatial information[J]. Journal of Advanced Transportation, 2016,50(8):2029-2040.
 [4] VLAHOIANNI E I, KARLAFTIS M G. Short-term traffic forecasting: where we are and where we are going[J]. Transportation Research

Part C: Emerging Technologies, 2014,43(1):3-19.
 [5] 赵恒辉,黄德启,曾蓉,等. 基于时空注意力 Bi-LSTM 模型的短时交通流预测[J]. 计算机仿真,2022,39(9):177-181,455.
 [6] TANG Jinjun, LIU Fang, ZOU Yajie, et al. An improved fuzzy neural network for traffic speed prediction considering periodic characteristic[J]. IEEE Transactions on Intelligent Transportation Systems, 2017,18(9):2340-2350.
 [7] YANCHONG C, DARONG H, LING Z. A short-term traffic flow prediction method based on wavelet analysis and neural network [C]//Proceedings of the 2016 Chinese Control and Decision Conference (CCDC). IEEE, 2016:7030-7034.
 [8] 宋旭东,任梦雪. 基于组合模型的短时交通流预测[J]. 计算机仿真,2022,39(7):156-160,406.
 [9] 肖宇,赵建有,叱干都,等. 基于 XGBoost 的短时出租车速度预测模型[J]. 交通信息与安全,2022,40(3):163-170.
 [10] WOLPERT D H, MACREADY W G. No free lunch theorems for optimization. [J]. IEEE Transactions on Evolutionary Computation, 1997,1(1):67-82.
 [11] PRUDÊNCIO R B C, LUDERMIR T B. Meta-learning approaches to selecting time series models[J]. Neurocomputing, 2004,61(1):121-137.
 [12] MA Shaohui, FILDES R. Retail sales forecasting with meta-learning [J]. European Journal of Operational Research, 2021,288(1):111-128.