

文章编号: 2095-2163(2023)03-0215-07

中图分类号: U491.2

文献标志码: A

# 基于模型融合的车辆跟驰模型研究

吴至锦, 曹从咏, 孔进, 陈鑫

(南京理工大学自动化学院, 南京 210094)

**摘要:** 为了提高车辆跟驰模型的精度与鲁棒性, 提出一种以 Stacking 为基础, 融合多个基学习器的车辆跟驰模型。采用网格搜索方法确定单一学习器的超参数, 考虑单一学习器的算法原理和预测性能, 选择 LGBM、SVR 和 KNN 作为基学习器, 通过对比分析的方式将 GBDT 模型作为元学习器, 以基学习器与元学习器为基础构建车辆跟驰模型。在 I-80 数据集上的实验结果表明: 构建的车辆跟驰模型相较于 IDM 模型在 MAE、MSE、RMSE 三项指标上分别降低了 0.305 6、3.230 4 和 1.031 5, 能够挖掘出数据之间的潜在关系, 弥补传统跟驰模型的不足。

**关键词:** 智能交通; 跟驰模型; 模型融合; Stacking; 机器学习

## Research on vehicle following model based on model fusion

WU Zhijin, CAO Congyong, KONG Jin, CHEN Xin

(School of Automation, Nanjing University of Science and Technology, Nanjing 210094, China)

**[Abstract]** In order to improve the accuracy and robustness of vehicle following model, a vehicle following model based on Stacking and integrating multiple base learners is proposed. The grid search method is used to determine the super parameters of a single learner. Considering the algorithm principle and prediction performance of a single learner, LGBM, SVR and KNN are selected as base learners. GBDT model is used as a meta learner through comparative analysis. The car following model is constructed based on basic learner and meta learner. The experimental results on the I-80 data set show that compared with the IDM model, the MAE, MSE and RMSE of the vehicle following model are reduced by 0.305 6, 3.230 4 and 1.031 5 respectively, which can dig out the potential relationship between the data and make up for the shortcomings of the traditional car following model.

**[Key words]** intelligent transportation; car following model; model fusion; Stacking; machine learning

## 0 引言

车辆跟驰 (Car Following, CF) 是微观驾驶中最基本的行为, 车辆跟驰模型是交通流理论的基础。20 世纪 60 年代初期, Chandler、Herman 和 Gazis 等学者提出并改进了 GM 模型, 开创了微观交通流理论的研究<sup>[1-2]</sup>。Michaels 等学者<sup>[3]</sup>、Kometani 等学者<sup>[4]</sup> 分别提出了心理-生理类模型和安全距离模型。2000 年, Treiber 等学者<sup>[5]</sup> 基于自动驾驶技术提出了智能驾驶员 (IDM) 模型。IDM 模型是目前较为完整、简单的无事故理论跟驰模型, 且该模型数值模拟得到的结果与实测数据一致性高。以上模型都可以统称为模型驱动跟驰模型, 其特点是有严谨的定义且具有严格的推导过程, 所运用的公式可解释性强<sup>[6]</sup>。

随着数据采集设备的更新与计算机技术的发展,

研究者可以获取高精度、大样本、瞬时车辆运动数据, 推动了数据驱动跟驰模型的快速发展。Kehtarnavaz 等学者<sup>[7]</sup> 首先通过车辆速度、车间距、方向盘角度与车头角度等数据构建了前馈神经网络的跟驰模型, 使得车辆能够进行自适应跟驰。随后, 国内有贾洪飞<sup>[8]</sup>、李德慧<sup>[9]</sup> 和徐学明<sup>[10]</sup> 等学者分别使用前馈神经网络、模糊神经网络与混合神经网络模拟单一车道的车辆跟驰行为。近年来, 随着美国 “Next Generation Simulation”<sup>[11]</sup> (NGSIM) 的交通仿真工程数据的开源, 越来越多的学者采用 NSIM 的数据开展研究。Wei 等学者<sup>[12]</sup> 基于 NSIM 车辆轨迹数据, 使用支持向量机建立了车辆跟驰模型。刘亚龙<sup>[13]</sup> 采用随机森林算法在 NSIM 数据集上建立车辆跟驰模型, 并与 Gipps 模型进行了对比。丁点点等学者<sup>[14]</sup> 使用线性组合的方式将 Gipps 模型和

**作者简介:** 吴至锦 (1998-), 男, 硕士研究生, 主要研究方向: 交通大数据; 曹从咏 (1964-), 男, 博士, 教授, 主要研究方向: 交通大数据、交通信息工程及控制。

**通讯作者:** 曹从咏 Email: jy108@njust.edu.cn

收稿日期: 2022-04-25

BP神经网络模型耦合起来并在NGSIM数据集上进行实验,结果表明耦合模型在预测速度时能够综合考虑真实性和安全性。

当前学者研究的数据驱动车辆跟驰模型涉及到神经网络、支持向量回归、随机森林、K近邻等多种机器学习方法,也有学者将机器学习方法与模型驱动跟驰模型结合起来,但鲜有学者将多种机器学习方法组合起来构建车辆跟驰模型。为此,本文基于NGSIM车辆跟驰实际场景,以多种机器学习方法为基础,使用Stacking的方式融合多种机器学习方法的优点,建立一个基于多学习器组合而成的车辆跟驰模型。

### 1 Stacking 模型融合原理

Stacking 是一种分层模型集成框架。以两层为例,第一层由多个基学习器组成,其输入为原始训练集,第二层的元学习器则是以第一层基学习器的输出作为特征加入训练集进行再训练,从而得到完整的Stacking模型。本文所使用的Stacking框架如图1所示。

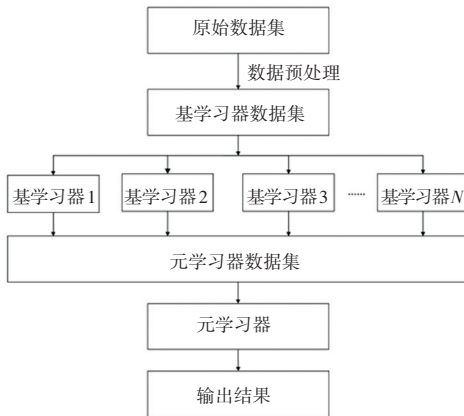


图1 Stacking 框架图  
Fig. 1 Stacking frame diagram

要得到Stacking模型,关键在于如何构造第二层的特征(下记为元特征,meta feature)。在构造元特征时,为了充分利用数据集并且使模型具有一定的泛化性能,Stacking对于原始训练集常常采用类似于K折交叉验证<sup>[15]</sup>的划分方法。

以5折划分为例,首先将原始训练集分为5份,分别记为 $f_1$ 、 $f_2$ 、 $f_3$ 、 $f_4$ 和 $f_5$ 。划分完成后,第一次使用 $f_1 \sim f_4$ 的数据来训练基模型,对 $f_5$ 以及测试集的数据进行预测, $f_5$ 的预测值 $f_5'$ 作为该基模型对 $f_5$ 生成的元特征,测试集的预测结果 $t_5$ 留作备用;第二次使用 $f_1 \sim f_3$ 与 $f_5$ 的数据来训练基模型,对 $f_4$ 及测试集的数据进行预测, $f_4$ 的预测值 $f_4'$ 作为该基模型对 $f_4$ 生成的元特征,测试集的预测结果 $t_4$ 留作备用。将上述操作继续重复3次,得到基模型对原始训练集生成的5个元特征 $f_1'$ 、 $f_2'$ 、 $f_3'$ 、 $f_4'$ 、 $f_5'$ 以及基模型对测试集预测得到的5个预测结果 $t_1$ 、 $t_2$ 、 $t_3$ 、 $t_4$ 、 $t_5$ ,将上述5个元特征垂直组合起来得到该基模型的训练结果,同时对于测试集,将上述5个预测值取平均值,作为测试集的输出进入到第二层模型中。对于其他基模型,也采用相同的方法生成元特征,从而构成用于第二层模型训练的完整元特征集。具体训练流程如图2所示。

### 2 数据来源与预处理

#### 2.1 数据来源

研究数据采用了NGSIM中研究人员于2005年4月13日下午4点至4点15分在I-80路段收集的 车辆轨迹数据。论文选取的研究路段如图3所示。

研究区域长度约为500 m,包括6条高速公路车道,还有一个入口匝道位于研究区域内。该车辆轨迹数据不仅提供了研究区域内每辆车辆的精确位置,也包括该车辆的车道位置和相对于其他车辆的位置,每条数据间隔时间为0.1 s。

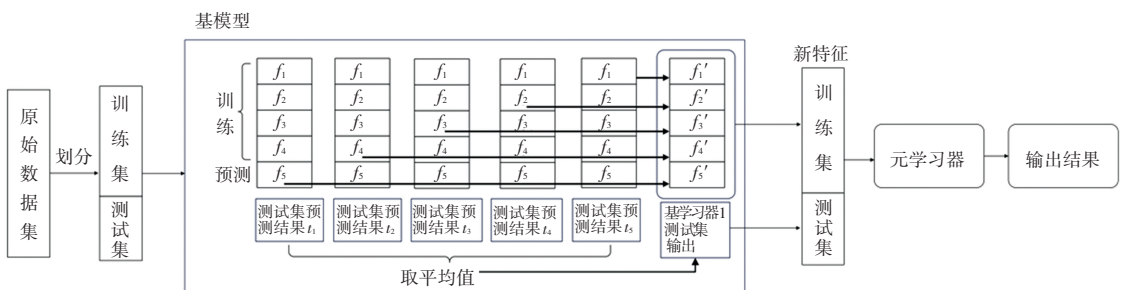


图2 Stacking 训练流程图  
Fig. 2 Stacking training flow chart

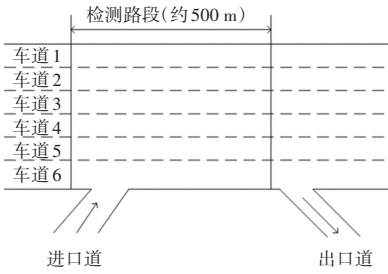


图3 I-80路段示意图

Fig. 3 Schematic diagram of section I-80

2.2 数据预处理

为保证用于模型训练的车辆轨迹数据包含充足

的、完整的有关车辆跟驰行为的信息,本文使用以下3条基本规则对上述数据集进行筛选,以挑选出合适的车辆跟驰数据:

(1)为了减少车型不同所导致的跟驰特性的差异,仅筛选出小型车的车辆轨迹数据。

(2)数据处理时,将主车与前导车作为一个跟驰对。

(3)满足跟驰时间大于等于26s、车头间距小于等于125m以及车头时距小于等于5s这3个条件的跟驰对为本文研究所需的有效跟驰对。经过筛选后的部分跟驰车辆数据见表1。

表1 跟驰车辆数据

Tab. 1 Car following data

跟驰车辆ID	前方车辆ID	车辆速度/ (m · s <sup>-1</sup> )	车辆加速度/ (m · s <sup>-2</sup> )	车头间距/ m	前方车辆速度/ (m · s <sup>-1</sup> )	前方车辆加速度/ (m · s <sup>-2</sup> )	下一时刻跟驰车辆速度/(m · s <sup>-1</sup> )
2	44	8.24	0.40	30.70	8.23	-0.05	10.18
7	21	6.96	1.73	32.66	5.61	0.00	7.16
11	1	3.82	0.00	13.46	3.60	-3.41	3.18
12	25	11.12	-0.16	18.59	7.17	0.00	11.42
17	2	2.21	-2.51	8.67	1.58	-0.01	1.11

原始数据集根据上述3条原则进行筛选后,共获得有效数据57928条,其中共有1192个有效跟驰对,训练集与测试集的比例为7:3。

3 车辆跟驰模型的建立

在建立基于Stacking的车辆跟驰模型前,首先要确定影响跟驰行为的关键变量。本文使用 $t$ 时刻跟驰车辆的速度 $v_n(t)$ 、加速度 $a_n(t)$ 、车头间距 $s_n(t)$ 、前导车的速度 $v_{n+1}(t)$ 、加速度 $a_{n+1}(t)$ 五个指标作为模型输入特征,以下一时刻跟驰车辆速度 $v_n(t + \Delta t)$  ( $\Delta t = 1$  s)作为模型输出。同时,为了解决特征之间量纲不同、数据处于不同数量级的问题,在建模时将5个特征进行归一化处理,归一化后所有的特征数据均在0~1之间。归一化处理中需用到的数学公式分别为:

单条车道上的车辆跟驰情形如图4所示,基于Stacking的车辆跟驰模型示意如图5所示。

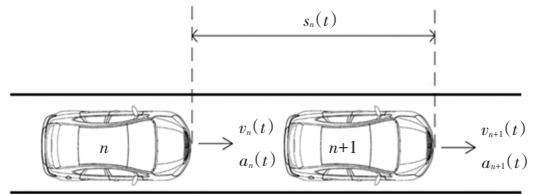


图4 车辆跟驰示意图

Fig. 4 Vehicle following diagram

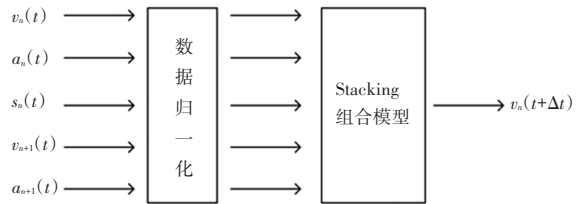


图5 基于Stacking的车辆跟驰模型

Fig. 5 Vehicle following model based on Stacking

4 基学习器和元学习器的选择

4.1 基学习器的选择

为构建基于多学习器融合的车辆跟驰模型,首先需要确定用于模型融合的基学习器。本文主要考虑了8种不同的学习器,包括以Boosting方法为核心的自适应提升算法(Adaptive boosting, AdaBoost)<sup>[16]</sup>、梯度提升决策树(Gradient Boosting Decision Tree,

$$X_{std} = \frac{X - X.\min(axis = 0)}{X.\max(axis = 0) - X.\min(axis = 0)} \quad (1)$$

$$X_{scaled} = X_{std} * (\max - \min) + \min \quad (2)$$

其中, $X_{std}$ 表示标准化结果; $X$ 表示需要归一化的数据; $X.\min(axis = 0)$ 表示每列中的最小值组成的行向量; $X.\max(axis = 0)$ 表示每列中的最大值组成的行向量。式(2)中,max表示要映射到的区间最大值,本文中取1,min表示要映射到的区间最小值,本文中取0。

GBDT)<sup>[17]</sup> 和由 GBDT 改进而来的 LightGBM (LGBM)<sup>[18]</sup>、极端梯度提升决策树 (Extreme Gradient Boosting, XGBoost)<sup>[19]</sup> 算法;以 Bagging 方法为核心的随机森林 (Random Forest, RF)<sup>[20]</sup> 算法;以统计学习 VC 维理论和结构风险最小化原理为基础的支持向量机回归 (SVR)<sup>[21]</sup> 算法;以具有前向结构的人工神经网络为基础的多层感知机 (MLP) 算法以及通过搜索

与当前观测值相似的历史数据进行匹配计算的 K 近邻 (K- Nearest Neighbor, KNN) 算法。

研究使用 Python 的 sklearn、lightgbm 库对上述学习器进行了实现,并且使用 sklearn 的网格搜索方法 (GridSearchCV) 确定上述 8 种学习器的参数。经网格搜索后,各模型在本文 2.2 节数据集上的超参数设置以及预测精度见表 2。

表 2 8 种学习器的训练结果

Tab. 2 Training results of eight basic learning devices

算法名称	参数	MSE	测试集 R <sup>2</sup>
LGBM	学习率=0.02,树的棵数=350,最大深度=7	1.182 8	0.895 8
AdaBoost	树的棵数=100	3.639 4	0.679 3
GBDT	学习率=0.05,最大深度=4,树的棵数=100	1.354 8	0.880 6
XGB	学习率=0.02,最大深度=4,树的棵数=400	1.358 2	0.880 3
MLP	激活函数='relu',隐含层数=1, 隐含层神经元个数=128	1.326 5	0.883 1
KNN	邻居个数=15	1.614 4	0.857 8
SVR	gamma = 0.3, 核函数='rbf'	1.204 9	0.893 8
RF	最大深度=8,树的棵数=50	1.419 9	0.874 8

根据单个学习器的预测结果和选择基学习器时‘好而不同’的原则<sup>[22]</sup>,本文首先排除了效果欠佳的 AdaBoost 模型。对剩余 7 种学习器的预测数据使用 MIC 最大信息系数<sup>[23]</sup>进行相关性分析,如图 6 所示。从图 6 中可以看出,学习器之间的 MIC 相关性系数普遍较高,造成这种现象的主要原因是各个学习器自身的学习能力较强,预测精度较高。其中,LGBM 和 XGB 模型是以 Boosting 为核心的算法且都从 GBDT 模型的基础上改进而来,因此在基学习器的选择时本文选取了这 3 种模型中精度最高的 LGBM 模型。

此外,KNN、SVR 和 LGBM 模型因为算法的实现机理有着很大的区别,从图 6 中也可以看出 KNN 模型与 SVR、LGBM 模型之间的相关性相较于其他模型之间的相关性偏低,因此本文最终确定 KNN、SVR 和 LGBM 三个模型作为模型融合的基学习器。

4.2 元学习器的选择

为了发挥不同模型的性能,本文首先考虑了在选择基学习器时未使用到的 MLP、XGB、GBDT、RF、AdaBoost 这五种模型。此外,为了探究不同元学习器对于模型精度的影响,还选取了 LASSO<sup>[24]</sup>、线性回归这 2 种多元线性回归模型。以上备选元学习器均在以 KNN、SVR 和 LGBM 为基模型的基础上进行对比验证,对比结果见表 3。

表 3 不同元学习器的预测结果

Tab. 3 Prediction results of different meta learners

元学习器	训练集 R <sup>2</sup>	测试集 R <sup>2</sup>	MAE	MSE	RMSE
GBDT	0.903 7	0.902 8	0.667 5	1.102 9	1.050 2
MLP	0.898 0	0.902 2	0.668 5	1.109 9	1.053 5
Liner	0.898 4	0.901 4	0.673 7	1.119 4	1.058 0
Lasso	0.898 4	0.901 3	0.674 1	1.120 7	1.058 6
XGB	0.923 0	0.897 3	0.680 3	1.165 3	1.079 5
RF	0.984 2	0.893 0	0.708 4	1.214 2	1.101 9
AdaBoost	0.781 7	0.814 9	0.918 9	2.100 9	1.449 4

由表 3 可以看出,由 GBDT 模型作为元学习器时拟合的精度最高。如果以 RF、XGB 作为元学习

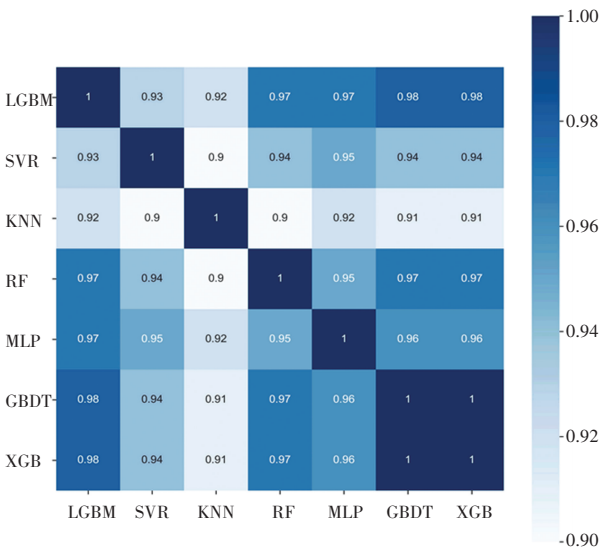


图 6 基学习器之间的 MIC 系数

Fig. 6 MIC coefficient between base learners

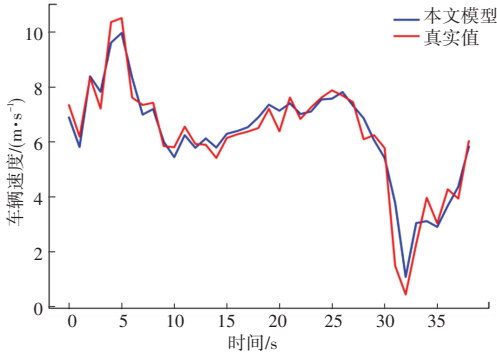


器,均产生了一定的过拟合现象,即相较于 GBDT、MLP 和多元线性回归模型,上述 2 种元学习器在训练集的拟合优度上升而测试集的拟合优度下降。此外,AdaBoost 模型在作为元学习器时仍展现出了较差的结果。为此,本文确定以 GBDT 模型作为元学习器。

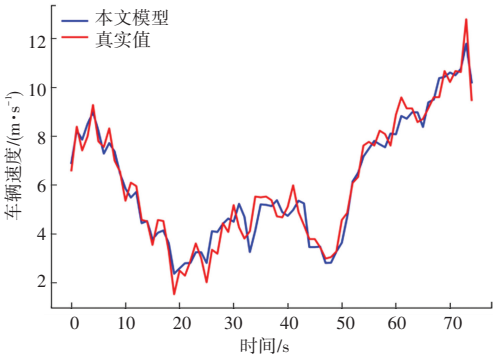
## 5 实验结果分析与讨论

### 5.1 实验结果分析

以 4.1 节与 4.2 节的研究为基础,将 LGBM、SVR 和 KNN 作为基学习器,GBDT 作为元学习器构建融合模型(下称本文模型),并在车辆跟驰数据上进行训练与仿真,得到本文模型的 MAE、MSE、RMSE 指标分别为 0.667 5、1.102 9、1.050 2,在训练集、测试集上的  $R^2$  分别为 0.903 7 和 0.902 8,计算耗时 280.6 s。部分仿真结果如图 7 所示,由仿真结果可以看出,本文模型能够较为准确地预判跟驰车辆的速度变化趋势。



(a) 2834 号跟驰车仿真结果



(b) 2799 号跟驰车仿真结果

图 7 本文模型仿真结果

Fig. 7 The simulation results of the combined model selected in this paper

### 5.2 本文模型与 IDM 模型的对比

为了证明本文所建跟驰模型的有效性,选择了 IDM 模型与本文模型进行结果仿真对比。IDM 模型的具体表达式为:

$$\alpha_n = a \left[ 1 - \left( \frac{v_n(t)}{\bar{v}} \right)^\sigma \left( \frac{S^*(v_n(t), \Delta v_n(t))}{\Delta s_n(t)} \right)^2 \right] \quad (3)$$

$$S^*(v_n(t), \Delta v_n(t)) = \bar{s} + v_n(t) T + \frac{v_n(t) \Delta v_n(t)}{2\sqrt{ab}} \quad (4)$$

其中,  $\alpha_n$  表示车辆加速度;  $v_n(t)$  表示跟驰车  $t$  时刻的车速;  $\Delta v_n(t)$  表示  $t$  时刻跟驰车与前导车的速度差;  $\Delta s_n(t)$  表示  $t$  时刻的车头间距;  $\bar{v}$  表示期望车速;  $\sigma$  表示加速度指数(本文取值为 4);  $a$  表示期望最大加速度;  $b$  表示舒适减速度;  $S^*$  为期望车头间距;  $\bar{s}$  表示静止安全距离;  $T$  表示安全车头时距。

为了避免 IDM 模型中的期望跟车距离  $S^*(v_n(t), \Delta v_n(t))$  出现负值<sup>[25]</sup>,并将模型的输出由加速度转化为下一时刻的车辆速度,将式(3)~(4)改写为:

$$v_n(t + \Delta t) = v_n(t) + \alpha_n \Delta t = v_n(t) + a \left[ 1 - \left( \frac{v_n(t)}{\bar{v}} \right)^\sigma - \left( \frac{S^*(v_n(t), \Delta v_n(t))}{\Delta s_n(t)} \right)^2 \right] \Delta t \quad (5)$$

$$S^*(v_n(t), \Delta v_n(t)) = \max \left\{ 0, v_n(t) T + \frac{v_n(t) \Delta v_n(t)}{2\sqrt{ab}} \right\} \quad (6)$$

本文取时间步长  $\Delta t$  为 1 s,安全车头时距  $T$  为 1.6 s<sup>[26]</sup>。对于其他参数,本文使用遗传算法在 2.2 节数据集的基础上进行参数标定,最终得到  $\bar{v} = 14.069 6$ ,  $a = 0.260 5$ ,  $b = 1.299 8$ ,  $\bar{s} = 4.773$ 。本文模型与 IDM 模型的各项误差指标见表 4。

表 4 本文模型与 IDM 模型对比评价

Tab. 4 Evaluation between the proposed model and IDM model

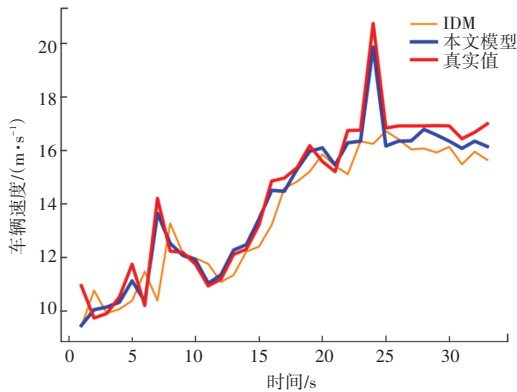
模型	MAE	MSE	RMSE
本文模型	0.667 5	1.102 9	1.050 2
IDM	0.973 1	4.333 3	2.081 7

由表 4 可知,本文模型相较于 IDM 模型预测效果更好,误差指标均有明显改善: MAE、MSE、RMSE 三项指标分别降低了 0.305 6、3.230 4 和 1.031 5,意味着本文模型的预测结果更加接近真实数据,能够适用于车辆跟驰模型。

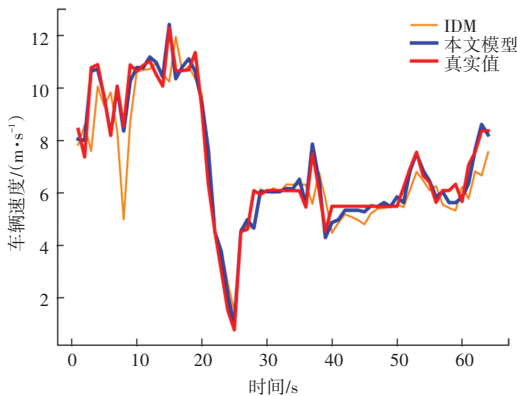
为了更加直观地将模型结果进行对比,选择了测试集中 2886 号与 1998 号跟驰车真实数据,绘制了仿真结果对比图,如图 8 所示。

由图 8(a)的结果可以看出, IDM 模型在第 7 s 和第 26 s 均没有很好地判别出车速的突然上升,而本文模型对于此种极端值的预测效果更好。这表明使用 Stacking 的方式可以使模型挖掘出数据之间的

潜在关系,弥补传统跟驰模型的不足,更好地对车辆的急加减速状况进行预判。综合图 8(a)和图 8(b),本文模型的仿真结果相较于 IDM 模型与真实数据更加贴合,具有更高的精度。



(a) 第 2886 号跟驰车仿真对比结果



(b) 第 1998 号跟驰车仿真对比结果

图 8 本文模型与 IDM 模型仿真结果对比

Fig. 8 Comparison between the proposed model and the simulation results of IDM model

## 6 结束语

研究采用 Stacking 模型融合的方法,选用 KNN、LGBM 和 SVR 模型作为基学习器,以 GBDT 模型作为元学习器构建车辆跟驰模型,通过对比证明了本文模型的优越性。本文总结了所选模型综合效果最优的 3 个原因:

(1) KNN、LGBM 和 SVR 三种基学习器在单一模型的情况下预测性能较强,可以提高模型的整体性能。

(2) 使用 Stacking 模型融合的方式可以进一步挖掘跟驰数据的潜在关系,弥补传统跟驰模型的不足。

(3) KNN、LGBM 和 SVR 在预测机理方面存在很大不同,采用差异性大的算法作为基学习器能够充分发挥每个算法的优势,使得融合模型有更好、更

稳健的预测性能。

本研究存在的不足有 2 点。一是在选择基学习器时采用了人工的方式,效率偏低;二是在对基学习器进行参数调优时选择了近似于穷举的网格搜索方法,增加了参数调优的工作量与时间。在未来的研究中,希望能在对基学习器进行调参和选择时结合优化算法,提升研究效率。

## 参考文献

- [1] BRACKSTONE M, MCDONALD M. Car-following: a historical review[J]. Transportation Research Part F Traffic Psychology & Behaviour, 1999, 2(4):181-196.
- [2] 王殿海,金盛. 车辆跟驰行为建模的回顾与展望[J]. 中国公路学报,2012,25(01):115-127.
- [3] MICHAELS R M. Perceptual factors in car following [C]// Proceedings of the Second International Symposium on the Theory of Traffic Flow. Paris:OECD,1963:44-59.
- [4] KOMETANI E, SASAKI T. Dynamic behaviour of traffic with a nonlinear spacing - speed relationship [C]//Proceeding of the Symposium on Theory of Traffic Flow. New York:Elsevier, 1959: 105-119.
- [5] TREIBER M, HENNECKE A, HELBING D. Congested traffic states in empirical observations and microscopic simulations[J]. Physical Review E, 2000, 62:1805-1824.
- [6] 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(05):102-113.
- [7] KEHTARNAVAZ N, GROSOWLD N, MILLER K, et al. A transportable neural - network approach to autonomous vehicle following[J]. Vehicular Technology IEEE Transactions on, 1998, 47(2):694-702.
- [8] 贾洪飞,隗志才,王晓原. 基于神经网络的车辆跟驰模型的建立[J]. 公路交通科技,2001(04):92-94.
- [9] 李德慧,刘小明,荣建,等. 基于模糊神经网络的车辆跟驰建模与仿真研究[J]. 北京工业大学学报,2007,33(04):398-401.
- [10] 徐学明,荣建,王丽. 混合神经网络跟驰模型的建立[J]. 公路交通科技,2007(03):130-132.
- [11] ALEXIADIS V, COLYAR J, HALKIAS J, et al. The next generation simulation program[J]. Ite Journal, 2004, 74(8):22-26.
- [12] WEI Dali, LIU Hongchao. Analysis of asymmetric driving behavior using a self - learning approach [J]. Transportation Research Part B Methodological, 2013, 47(1):1-14.
- [13] 刘亚龙. 基于随机森林的车辆跟驰行为模型[J]. 山东交通学院学报,2018,26(04):45-50.
- [14] 丁点点,孙磊,陈松. 机器学习——动力学耦合车辆跟驰模型[J]. 交通运输系统工程与信息,2017,17(06):33-39.
- [15] FUSHIKI T. Estimation of prediction error by using K-fold cross-validation[J]. Statistics and Computing, 2011, 21(2):137-146.
- [16] FREUND Y, SCHAPIRE R E. A desicion-theoretic generalization of on-line learning and an application to boosting[J]. Journal of Computer and System Sciences, 1997, 55:119-139.
- [17] FRIEDMAN J H. Greedy function approximation: A gradient boosting machine[J]. Annals of Statistics, 2001, 29(5):1189-1232.

(下转第 225 页)