

基于 GRU-SVR 的短时交通流量预测研究

潘伟靖^{1,2}, 陈德旺^{1,2}

(1. 福州大学 数学与计算机科学学院, 福建 福州 350108;
2. 福州大学-星云股份智慧新能源研究中心, 福建 福州 350108)

摘要:短时交通流量预测为智能交通系统(ITS)的研究方向之一。现有相关研究中所提及的基于深度学习的方法,需要较高的计算复杂度或模型的回归预测能力存在一定的不足。因此,提出一种将门循环单元(gated recurrent unit, GRU)与支持向量回归(support vector regression, SVR)相结合的模型。该模型一方面借助深度学习模型的强大能力进行特征提取工作,并相较于短时记忆网络(long short term memory, LSTM)降低了一定的计算量,同时,又以支持向量回归模型来增强整个模型的回归预测能力。文中基于 Keras 所提供的 python 库,完成实验设计及开发,根据实验结果对模型进行逐步调整,选择最优模型,并在 PeMS 数据集上对模型的泛化能力以及抗噪能力进行了测试。实验结果表明,与 SVR 以及 GRU 模型相比较,GRU-SVR 模型预测精度分别提升了 4% 和 1.6%,并且该模型具有一定的泛化及抗噪能力。

关键词:门限循环单元;支持向量回归;长短期记忆网络;深度学习;短时交通流量

中图分类号:TP31

文献标识码:A

文章编号:1673-629X(2019)10-0011-04

doi:10.3969/j.issn.1673-629X.2019.10.003

Research on Short-term Traffic Flow Prediction Based on GRU-SVR

PAN Wei-jing^{1,2}, CHEN De-wang^{1,2}

(1. School of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China;
2. Nebula Intelligent New Energy Research Center of Fuzhou University, Fuzhou 350108, China)

Abstract: Short-term traffic flow forecasting is one of the research directions of intelligent transportation systems (ITS). The deep learning-based approach mentioned in the existing related research requires high computational complexity or the regression prediction of the model is insufficient. Therefore, we propose a model that combines gated recurrent unit (GRU) with support vector regression (SVR). On the one hand, this model makes use of the powerful ability of deep learning model to carry out feature extraction, and reduces the computation amount compared with long short term memory (LSTM). On the other hand, support vector regression model is used to enhance the regression prediction of the whole model. The experimental design and development is completed based on the python library provided by Keras. And the optimal model is selected according to the experimental results, and the generalization and anti-noise of the model are tested on the PeMS dataset. The experiment shows that the GRU-SVR model has improved prediction accuracy by 4% and 1.6% in contrast with SVR and GRU. And the model has certain generalization and anti-noise ability.

Key words: gated recurrent unit; support vector regression; long short term memory; deep learning; short-term traffic flow

0 引言

短时交通流量预测研究,为智能交通系统中的一个重要研究方向。该类研究可以为相关部门提供较为准确的城市道路流量信息,进而为相关部门对于路面车流的引导以及预警信息的发布提供相关决策依据,为市民的出行等提供了更为可靠的建议。因此,该类研究也具有一定的现实意义。

在短时交通流量预测的问题上,很多学者做了相

关研究,如宫晓燕等使用的基于非参数回归的方法^[1]以及张一基于卡尔曼滤波算法的研究^[2]等一系列基于数学模型进行的研究。杨兆升等采用支持向量机^[3]等机器学习方法进行预测。有部分学者采用一些混合模型,期望多种方法的互补,使得其模型的最终预测达到更好的效果,如采用 SVM 与卡尔曼滤波法相结合的研究^[4]、将 K-means 与极限学习机相结合的方法^[5]以及采用小波分析法与 ARMA 模型相结合的方法^[6]

收稿日期:2018-11-16

修回日期:2019-03-19

网络出版时间:2019-06-26

基金项目:国家重点研发计划课题(2018YFB0104403);福州大学校企合作项目(01001707,00101522)

作者简介:潘伟靖(1996-),男,硕士研究生,研究方向为机器学习、大数据等;陈德旺,教授,博导,研究方向为大数据、智能交通系统等。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20190626.0829.036.html>

等,在短时交通流量预测方向,取得了明显的进展。

随着深度学习及硬件设备等的蓬勃发展,不少学者尝试构建深层次神经网络,希望网络模型在预测结果精确度和可靠性上更高。其中对短时交通流量预测这类时间序列预测问题而言,最常用的便是 RNN。

RNN 是一种区别于传统前馈神经网络的新型网络,其设计思想是使神经网络具有一定的“记忆”功能,从而使训练后的 RNN 模型可以根据已有的若干个状态的值来对下一时刻 S 的值进行估测,正是因为这种“记忆”的特性,为时间序列预测、字符序列等问题提供了一个新的解决方案。但是,在训练标准 RNN 试图获取数据在一段较长时间内的特征时,RNN 通常会出现梯度消失或梯度爆炸的情况^[7]。于是有研究人员尝试将 RNN 的隐藏层基本单位由一个简单神经元节点改变为一个拥有记忆性的模块^[8]。这种对标准 RNN 的更改中,最为著名的为 Hochreiter 和 Schmidhuber 提出的长短时记忆网络(long short term memory, LSTM)^[9]以及 Cho 等提出的门循环单元(gated recurrent unit,GRU)^[10]。这两种改进型的 RNN,使得对一段相对较长时间内的特征捕获问题有了长足的进步。目前,RNN 模型广泛应用在语音识别^[11]、自然语言处理^[12]、机器翻译^[13]、语义分析^[14]等方面。

有学者将 LSTM 运用于短时交通流量预测^[15],也有学者对 LSTM 与 GRU 两种模型在短时交通流量预测方面的性能做了比较^[16],也有学者将 LSTM 与 SVR 进行结合,得到了相对较好的结果^[8]。由于 GRU 相对 LSTM 通常收敛较快^[16],计算量也就会相对减少。在结构上只包括更新门和重置门。其中更新门会对当前时刻保存传来的状态信息多少做出决策,重置门则正好相反,是对当前时刻丢弃或忽略传来的状态信息的多少做出判定。在 LSTM 与 SVR 相结合^[8]的方法的启发下,文中提出一种将 GRU 与 SVR 相结合的模型用于短时交通流量预测,以期取得更高的预测精度和更快的运算速度。该模型使用门循环单元(GRU)对数据做特征提取,再将其输入到模型顶层的支持向量机回归(SVR)模型中以得到模型最终的预测输出。

1 模型设计

1.1 模型设计

提出的模型结构如图 1 所示。

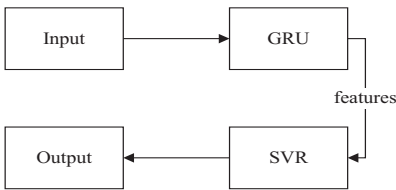


图 1 GRU-SVR 模型框图

假设在模型训练时一次输入模型的数据长度为 num,数据采样时间间隔设为 Δt ,即一次输入模型的一组数据之间的时间间隔为 $\text{num} * \Delta t$,且训练集整体表示为 $D = (x_1, x_2, \dots, x_n)$,其中 x_i 表示每个数据采样时间间隔内的交通流量数据。将训练集数据转化为向量 V ,表示为 $V = (V_1, V_2, \dots, V_m)$,其中 $V_i = (x_i, x_2, \dots, x_{i+\text{num}})$ 。之后对 V_i 做切片,使得其前 num 个数据为模型的一次输入,最后一个为其对应的结果。在这之后进行汇总,形成真正的训练数据。对测试集数据也做相同的处理,得到测试数据。

将训练数据作为图 1 中的 Input 输入模型,训练 GRU 网络模型使之学习到相关参数。之后将数据集输入已训练好的 GRU 模型,获取到由 GRU 模型输出的特征向量,组成新的训练集与测试集。输入到 SVR 模型进行训练,得到最终的预测信息。相应流程如下:

Step1:对获取到的数据进行规范化处理;

Step2:以处理后的数据训练 GRU 网络模型,对输入数据进行学习;

Step3:将 GRU 网络输出的特征向量作为 SVR 模型的输入,并对 SVR 模型进行训练;

Step4:使用 SVR 模型进行预测,并将模型输出做反规范化得到预测结果。

1.2 具体设计

在 GRU 网络设计中,首先固定输入数据的长度进行设计。这里以 num 为 12 来进行设计。经反复迭代测试,在兼顾模型的计算量与最终性能的情况下,选择使模型在测试集上 MAPE 值达到最小的模型,并不断调整其余超参数,使得模型的平均绝对百分比误差(mean absolute percentage error,MAPE)以及平均绝对误差(mean absolute error,MAE)等指标逐步降低。GRU 网络模型如图 2 所示。

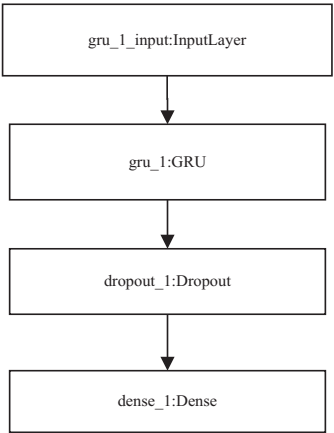


图 2 GRU 特征提取模型

在图 2 中,使用只有一层 GRU 网络层的网络结构,其中输入层单元数为 12,隐含层的单元数为 135,激活函数选择 sigmoid,初始化函数选择 he_uniform。

之后,Dropout 层随机断开 50% 的单元,再经过全连接层进行输出。Epochs 为 300。

在 GRU 网络结构确定之后,需要对 SVR 模型进行设计,在反复迭代测试之后,最终采用核函数为径向基(RBF),惩罚因子 $C = 1$, gamma 为 0.3, epsilon 为 0.001 的 SVR 模型。

2 实验数据及评价指标

2.1 实验数据

实验采用的数据来自加利福尼亚交通局公路绩效监控系统(Caltrans performance measurement system, PeMS),使用的是加利福尼亚州萨克拉门托市 I80-E 检测到的采样间隔为 5 分钟的流量数据,时间为 2018 年 4 月到 2018 年 7 月,且使用的为工作日的总车流量。选择将 4 月的工作日数据作为训练集对模型进行训练,5 月的工作日数据作为测试集。

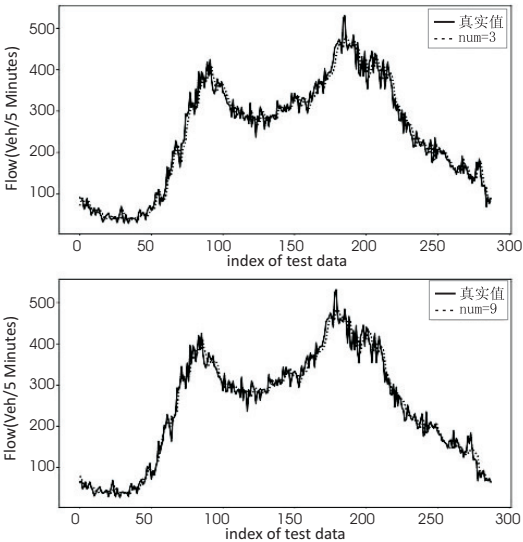


图 3 不同 num 的模型对测试集数据的预测结果

使用上文所述指标进行评价,其结果如表 1 所示。

表 1 不同数据输入长度模型的预测结果评价

num	MAPE/%	MAE
3	9.33	16.96
6	9.16	16.64
9	8.90	16.14
12	8.76	15.84

如图 3 所示,在所有的实验中,单次数据输入长度为 12 的模型所预测出的曲线,与真实流量的曲线始终较为贴合,即误差始终在一个相对较小的范围内。同时,从表 1 中可以看到,在不同实验中,以 num 为 12 训练的模型,使得模型的预测误差值达到最低,在其之前的几个不同实验中,MAPE 与 MAE 指标值均高于 num 为 12 的模型,但是都相差不大。所以仍然选择数据输

2.2 评价指标

采取的评价标准为平均绝对误差(MAE)以及平均绝对百分比误差(MAPE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |f_i - \hat{f}_i|$$
 (1)

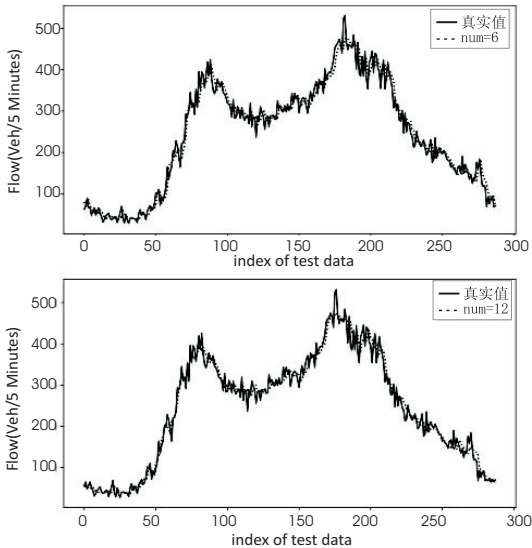
$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|f_i - \hat{f}_i|}{f_i} \times 100\%$$
 (2)

其中, f_i 为数据集中的实际值; \hat{f}_i 为模型对当前时刻的预测输出。

3 实验结果及分析

3.1 实验分析

在确定网络模型及其所配置的使 MAPE、MAE 等评价标准最低的超参数后,设计采用不同的单次数据输入长度进行实验。实验的数据输入长度包括 3,6,9,12。模型的预测结果如图 3 所示。



入长度为 12 的模型进行预测,达到最佳的结果。

3.2 模型性能测试

为了测试该模型的泛化性,使用 4 月数据训练所得模型,分别对获取到的 6、7 月数据进行测试,其结果如表 2 所示。

表 2 不同测试集上的性能比较

数据	MAPE/%	MAE
6 月	8.65	16.60
7 月	9.02	17.17

由表 2 可知,使用 4 月份数据作为训练集所训练得到的模型在 6、7 月份的数据集也有与 5 月份数据集上的相似表现,即使用 6、7 月数据作为测试集,其 MAE、MAPE 值与 5 月份数据作为测试集的对应评价指标浮动相对较小,说明该模型在非连续自然月的数

数据集上具有一定的泛化能力。

在这之后,对模型的抗噪能力也进行了测试。分别为 5 月份数据集加入三组具有不同均值、方差的高斯噪声以检测模型在不同情况下的抗噪能力。约定以“均值-标准差”的形式来表示每次测试的均值与标准差,其结果如表 3 所示。

表 3 加入不同高斯白噪声的性能比较

数据	MAPE/%	MAE
1-0.1	8.72	15.85
1-0.5	8.74	15.86
5-0.5	8.47	16.01

由表 3 可知,在加入高斯白噪声之后,模型的 MAPE、MAE 评价指标与原始测试集相比并没有产生过多的浮动。所以,可以说明该模型具有相对较好的抗噪性。

3.3 模型对比

为了验证该模型的性能,使用如上文所述的数据集进行训练及预测,即使用 4 月的数据进行训练,5 月的数据进行预测,且分别与相同参数的 GRU、SVR 模型的性能进行对比,其具体结果如图 4 所示。

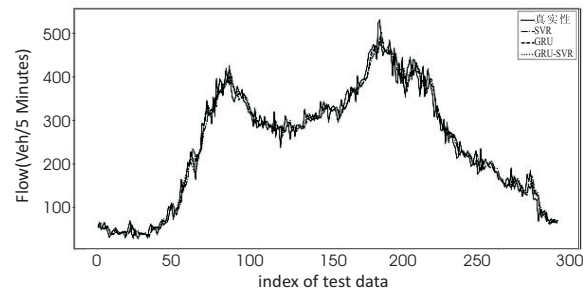


图 4 不同模型的预测结果对比

其预测的误差如表 4 所示。

表 4 不同模型的预测误差

模型	MAPE/%	MAE
SVR	9.16	16.51
GRU	8.90	16.11
GRU-SVR	8.76	15.84

由图 4 及表 4 可以看出,所有模型的预测曲线均可以较好地模拟出真实曲线数据,均较为贴合,相对误差均较小。但是,文中采用的模型相比于 SVR、GRU,可以取得最低的 MAPE 以及 MAE 的值,说明该模型优于 SVR 和 GRU 模型,且 GRU-SVR 模型最终的 r^2 值为 0.97,是一种可靠的预测模型。

4 结束语

借助深度学习模型强大的特征提取能力以及支持向量机回归的预测能力,以结合门循环单元 (GRU) 与

支持向量机回归 (SVR) 两种模型的混合模型对单一传感器的交通流量进行了相对较好的预测。但是,在大数据背景下,如何将机器学习与大数据相结合,仍是一个有前景的研究问题。对多个传感器在大数据背景下的短时交通流量预测问题,是下一步的研究内容。

参考文献:

[1] 宫晓燕,汤淑明. 基于非参数回归的短时交通流量预测与事件检测综合算法[J]. 中国公路学报,2003,16(1):82-86.

[2] 张 一. 基于卡尔曼滤波的短时交通流量预测模型研究[D]. 沈阳:沈阳工业大学,2014.

[3] 杨兆升,王 媛,管 青. 基于支持向量机方法的短时交通流量预测方法[J]. 吉林大学学报:工学版,2006,36(6):881-884.

[4] 朱征宇,刘 琳,崔 明. 一种结合 SVM 与卡尔曼滤波的短时交通流预测模型[J]. 计算机科学,2013,40(10):248-251.

[5] 芮兰兰,李钦铭. 基于组合模型的短时交通流量预测算法[J]. 电子与信息学报,2016,38(5):1227-1233.

[6] 杨芳明,朱顺应. 基于小波的短时交通流预测[J]. 重庆交通学院学报,2006,25(3):99-102.

[7] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE Transactions on Neural Networks,1994,5(2):157-166.

[8] 乔松林,孙仁诚,刘 吉. 基于深度学习的短时交通流量预测[J]. 青岛大学学报:自然科学版,2017,30(4):65-69.

[9] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation,1997,9(8):1735-1780.

[10] CHO K, VAN MERRIËNBOER B, GULCEHRE C, et al. Learning phrase representations using RNN encoder-decoder for statistical machine translation[C]//Conference on empirical methods in natural language processing. Stroudsburg, PA:Association for Computational Linguistics,2014:1724-1734.

[11] 朱小燕,王 昱,徐 伟. 基于循环神经网络的语音识别模型[J]. 计算机学报,2001,24(2):213-218.

[12] 朱晶晶,韩立新. 基于 RNN 句子编码器的聊天机器人[J]. 计算机与现代化,2018(1):32-35.

[13] 李婧萱. 基于深度神经网络的统计机器翻译模型研究[D]. 哈尔滨:哈尔滨工业大学,2015.

[14] 胡新辰. 基于 LSTM 的语义关系分类研究[D]. 哈尔滨:哈尔滨工业大学,2015.

[15] 王祥雪,许伦辉. 基于深度学习的短时交通流预测研究[J]. 交通运输系统工程与信息,2018,18(1):81-88.

[16] FU Rui, ZHANG Zuo, LI Li. Using LSTM and GRU neural network methods for traffic flow prediction[C]//Youth academic conference of chinese association of automation. Wuhan, China:IEEE,2017:324-328.