



# 基于改进的 LSTM 算法的时间序列流量预测

## 摘要

时间序列流量的预测问题是近年来机器学习的一个热点问题,通过改变长短期记忆网络(LSTM)层数、网络层神经元的个数、网络层之间的连接方式,特殊网络层的应用等网络结构以及优化器和损失函数的选择可以极大地提高预测的精度.本文提出多层 LSTM 算法,该算法是在传统 LSTM 算法上进行改进的单一模型,模型设计的复杂度低,可以提高机器学习的效率.模型采用一个输入层、5 个隐藏层、1 个输出层,同时包含 1 个全连接层和 1 个 Dropout 层,Dropout 层的作用是防止机器学习过拟合.选择 adam 为模型优化器、mse 为模型损失函数、relu 作为模型的激活函数.实验结果表明,与传统模型相比,该模型具有较好的泛化能力.

## 关键词

时间序列预测;长短期记忆网络;预测精度;泛化能力

中图分类号 TP393.06;TP18

文献标志码 A

收稿日期 2021-09-11

资助项目 国家自然科学基金(61803275);辽宁省“兴辽英才计划”项目(XLYC1907044);辽宁省自然科学基金(2020-MS-218);辽宁省教育厅重点项目(lnzd202007)

## 作者简介

郭佳丽,女,硕士生,研究方向为机器学习、优化算法.m18846761923@163.com

邢双云(通信作者),女,博士,教授,研究方向为随机控制、采样控制、智能优化控制、随机奇异系统理论等.xsy\_angel25@163.com

1 沈阳建筑大学 理学院,沈阳,110168

2 沈阳建筑大学 教务处,沈阳,110168

## 0 引言

时间序列预测<sup>[1-2]</sup>是机器学习的一个重要领域,其典型问题有交通流的预测<sup>[3]</sup>、旅游人数的预测、天气的预测、网络流量的预测<sup>[4]</sup>等.智能交通系统<sup>[5-7]</sup>(ITS)被视为缓减交通压力、减少汽车污染物的排放、降低各种交通事故的有效手段.交通流预测性能的提高是当下智能交通系统亟需解决的问题.交通流的预测可分为长时交通流预测和短时交通流预测.长时交通流预测以年、月、日为单位,主要用于道路规划与设计.短时交通流以 min 为单位,大多小于 15 min,主要用于交通控制或作为引导系统的依据.针对交通流量的不确定性、非线性以及规律性弱等特点,已有不同的组合模型来进行交通流量的预测.例如:Du 等<sup>[8]</sup>提出深度不规则卷积残差 LSTM(Long Short-Term Memory,长短期记忆)网络模型,集成了不规则卷积住宅网络和 LSTM 单元的深度学习框架来学习时空特征用于城市交通客流预测;武佳琪等<sup>[9]</sup>提出了 LSTM 和 BP 算法组合而成的预测道路堵塞时间的智能模型;李磊等<sup>[10]</sup>提出分时段预测交通流的思想,设计了 CNN-LSTM 的组合预测模型;Lu 等<sup>[11]</sup>提出了 ARIMA-LSTM 组合的交通流量预测算法,利用滚动回归 ARIMA 模型获取交通数据的线性回归特征,再利用反向传播对 LSTM 网络进行训练,获取交通数据的非线性特征.网络规模的日益增大,应用需求也变得多种多样,网络拥塞、故障等问题接踵而来,因此网络流量预测模型<sup>[12]</sup>的建立至关重要.黎佳玥等<sup>[13]</sup>设计了一种 LSTM 组合预测分类算法对异常流量进行预测,通过训练遭受典型攻击事件的大规模网络流量公共数据集,用以预测来自网络流量信号的典型异常攻击事件出现概率;王海宁等<sup>[14]</sup>提出了一个基于 LSTM 神经网络与传统神经网络的网络流量预测模型,与传统的机器学习的流量预测方法相比,该模型在网络流量预测中具有适用性和较高的精确度.

上述模型均是通过与 LSTM 人工神经网络组合生成组合模型对时间序列流量进行预测的,但组合模型涉及多个神经网络,所以复杂度高、机器学习效率较低.神经网络的作用是把输入数据的特征抽象到另一个维度空间,来展现其更抽象化的特征,多个网络层其实是对输入特征多层次的抽象,最终目的是为了更线性地划分不同类型的数据,所以本文提出多层 LSTM 算法的单一模型.经由实验验证,与文献<sup>[15]</sup>传统的 LSTM 相比,多层 LSTM 算法具有更优的泛化能力,预测效果更好.

### 1 长短期记忆网络(LSTM)

LSTM 是一种改进的循环神经网络 (Recurrent Neural Network, RNN), 它解决了 RNN 中的梯度消失和长期依赖的问题. LSTM 是在 RNN 的基础上进化而来的. 在标准的 RNN 网络中, 循环模块只有一个 tanh 层, 如图 1 所示.

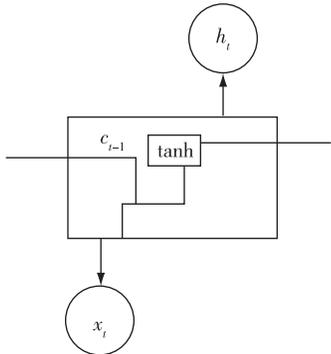


图 1 RNN 中的重复模块  
Fig. 1 Repeated module in RNN

LSTM 也有类似的链结构, 但是循环模块的结构不同. 它有 4 个神经网络层, 以一种特殊的方式相互作用. “门”是 LSTM 的重要组成部分, 门的“开”和“关”决定了历史信息的保留和删除, 门的工作原理如图 2 所示.

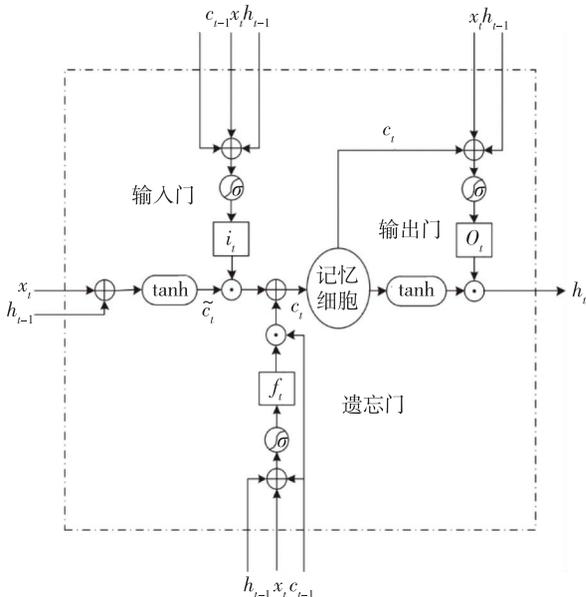


图 2 LSTM 基本工作原理  
Fig. 2 Basic working principle of LSTM

遗忘门  $f_t$  决定是否将上一次单元状态  $c_{t-1}$  保留给当前时间  $c_t$  以及哪些信息要丢弃. 遗忘门的计算

公式如下:

$$f_t = \sigma(\mathbf{w}_f \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_f) . \quad (1)$$

输入门  $i_t$  确定当前时间输入  $x_t$  保留到单元状态  $c_t$  的信息的多少, 以过滤掉不相关的信息. 输入门的计算公式如下:

$$i_t = \sigma(\mathbf{w}_i \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_i) . \quad (2)$$

输出门  $o_t$  要控制状态  $c_t$  输出多少信息到 LSTM 的当前输出值  $h_t$ , 并控制长期内存对当前输入的影响. 输出门的计算公式如下:

$$o_t = \sigma(\mathbf{w}_o \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_o) . \quad (3)$$

输出结果  $h_t$  的计算公式如下:

$$\mathbf{h}_t = o_t * \tanh(\mathbf{c}_t) , \quad (4)$$

其中 tanh 和  $\sigma$  是两个常用的激活函数, 计算公式为

$$\sigma(x) = \frac{1}{1 + \sqrt{e}^x} , \quad (5)$$

$$\tanh(x) = \frac{\sinh x}{\cosh x} . \quad (6)$$

在上述公式中,  $\mathbf{w}$  是隐藏层的权值矩阵,  $\mathbf{b}$  是偏移量.

### 2 改进的 LSTM 算法

#### 2.1 模型搭建

首先, 经过大量的实验, 我们发现在神经网络的使用过程中, 神经网络的层数、神经元的个数以及各层之间的连接方式对流量预测的结果影响很大, 因此, 本文设计了多层 LSTM 人工神经网络. 综合考虑时间序列的预测精度和降低模型时间复杂度的问题, 预测精度与其他组合模型基本相同, 但由于简化了模型设计的复杂度, 使得机器运行速度更快, 机器学习效率更高. 其次, 网络层数的增加会使机器学习面临过拟合的问题, 因此本文模型在网络层中插入了 Dropout 层, Dropout 可以随机断开网络层之间的连接, 进而减少训练时实际参与的模型的参数量, 从而减少模型的实际容量, 可以防止过拟合. 经实验验证, 选取 Dropout 的值为 0.4 时可以达到最优效果. 改进后的 LSTM 人工神经网络结构 (多层 LSTM 模型 1) 如图 3 所示, 该模型选择 adam 为模型优化器、mlse 为模型损失函数、relu 为模型激活函数.

为了验证网络层对模型预测能力的影响以及本文搭建的多层 LSTM 人工神经网络的有效性, 同时搭建了三层 LSTM 人工神经网络 (多层 LSTM 模型 2), 如图 4 所示.

模型采用滑动窗口的方式划分数据集和验证

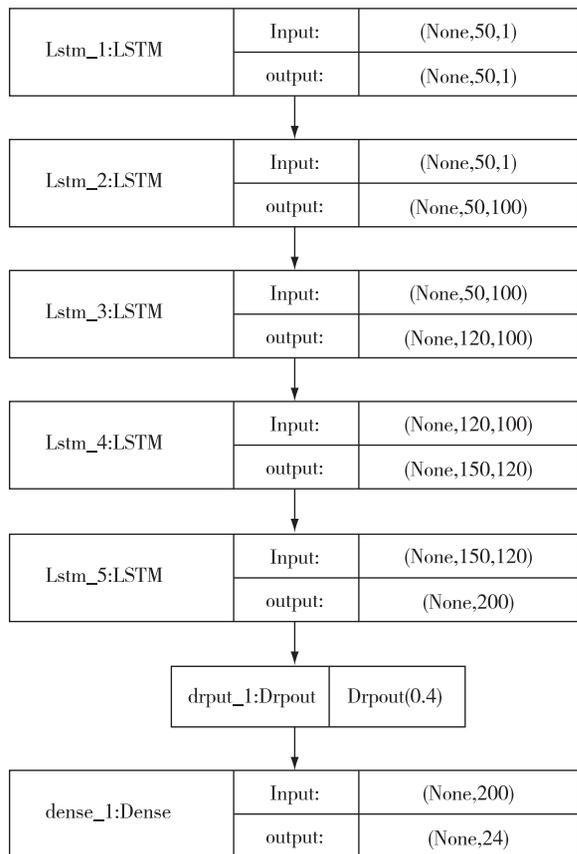


图3 多层 LSTM 模型

Fig. 3 Multilayer LSTM model

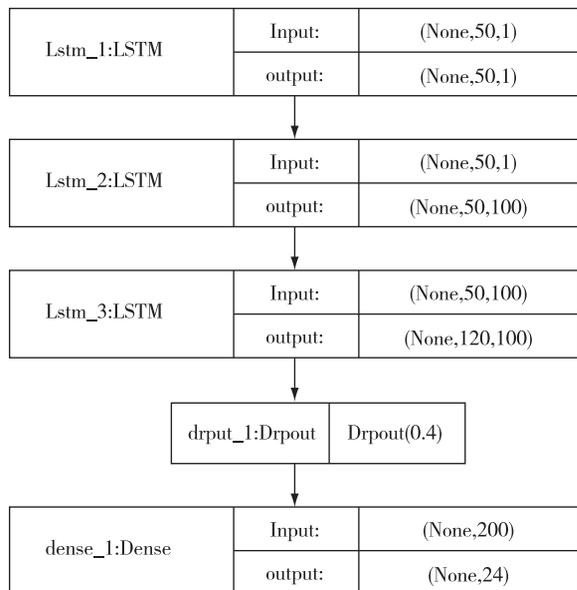


图4 三层 LSTM 模型

Fig. 4 Three-layer LSTM model

集,利用随机梯度下降法以及误差的反向传播原理不断地减小损失函数,自动更新权重,通过合理次数

的迭代训练,使模型拥有相对更优的学习能力。

本文选用 Google 开源的基于数据流图的机器学习框架 tensorflow,以及 Keras 作为深度学习库.在 16 GB 内存和 i5-11300H 的机器上完成训练过程。

### 2.2 性能指标

本文采用均方根误差 (RMSE, 量值记为  $\eta_{RMSE}$ )、平均绝对误差 (MAE, 量值记为  $\eta_{MAE}$ ) 和平均绝对百分比误差 (MAPE, 量值记为  $\eta_{MAPE}$ ) 三个误差指标来评价模型的预测能力:

$$\eta_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n |f_i - \hat{f}_i|^2}, \quad (7)$$

$$\eta_{MAE} = \frac{1}{n} \sum_{i=1}^n |f_i - \hat{f}_i|, \quad (8)$$

$$\eta_{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|f_i - \hat{f}_i|}{f_i}, \quad (9)$$

其中,  $f_i$  和  $\hat{f}_i$  分别表示实际值和模型的预测值。

## 3 验证和模拟

### 3.1 数据来源

数据来自英国政府公布的公共高速公路网络数据库 (<http://tris.highwaysengland.co.uk/detail/traffic-flowdata>).数据集网站包含英国 M、A 级所有高速公路,数据非常全面,以 15 min 为间隔采集,由传感器自动收集,或由操作员手动收集.本文提取的数据为 M25 高速希思罗机场附近的 2019 年 8 月的 3 个站点的数据集.每个数据集包含 2 976 个数据.图 5 为其中一个站点的原始数据,可以看出该数据有明显的周期性且平稳,可以作为时间序列流量预测的数据集。

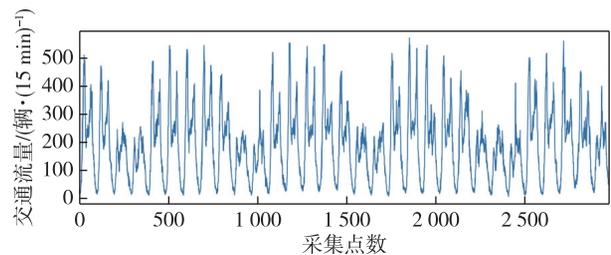


图5 原始序列

Fig. 5 Original sequence

### 3.2 验证结果及分析

为了验证多层 LSTM 算法在时间序列预测方面的作用,本文采用 3 个数据集对传统 LSTM 算法与改进不同程度的多层 LSTM 算法进行训练和测试.用

RMSE、MAE、MAPE 三个误差评价指标对改进的 LSTM 算法的有效性进行验证,结果如表 1 所示。

表 1 模型改进前后预测结果比较  
Table 1 Comparison of indicators before and after model improvement

模型	指标	站点 1	站点 2	站点 3
传统 LSTM 模型	RMSE	44.185 9	135.406 6	73.533 4
	MAE	30.644 5	104.795 4	59.772 2
	MAPE	19.731 9	13.315 5	19.908 1
多层 LSTM 模型 2	RMSE	31.193 8	114.891 9	60.378 9
	MAE	24.946 6	101.187 5	52.818 3
	MAPE	18.861 7	10.906 0	20.104 9
多层 LSTM 模型 1	RMSE	23.767 7	110.513 5	36.019 0
	MAE	17.430 9	90.478 6	28.158 8
	MAPE	9.251 1	10.196 5	10.401 1

由表 1 可以看出,多层 LSTM 模型 1 的各项误差远小于传统的单层 LSTM 算法。由图 6 可知:改进后的模型预测值更接近实际值,表明改进后模型的预测精度更高。

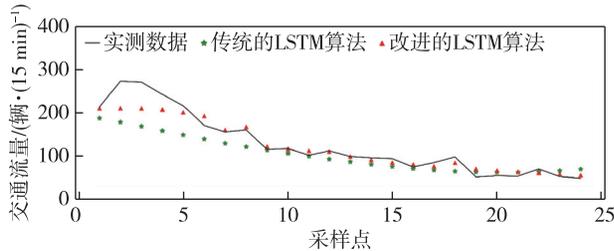


图 6 预测值与真实值结果对比

Fig. 6 Comparison between predicted values and real values

## 4 结论

本文考虑到 LSTM 的网络层数、神经元的个数、网络层之间的连接方式等网络结构以及优化器和损失函数的作用,提出了多层 LSTM 算法的时间序列流量的预测模型。本文模型综合考虑了时间序列预测精度和模型设计的复杂度。与传统的 LSTM 神经网络相比,本文模型具有更好的预测效果。本文仅采用历史流量数据来模拟和预测未来的流量数据,后期可以通过增加输入数据的维度来提高时间序列流量的预测精度。

## 参考文献

### References

[ 1 ] Xu M L, Han M, Chen C L P, et al. Recurrent broad

郭佳丽,等.基于改进的 LSTM 算法的时间序列流量预测.  
GUO Jiali, et al. Prediction of time series traffic based on improved LSTM algorithm.

- learning systems for time series prediction [ J ]. IEEE Transactions on Cybernetics, 2020, 50 ( 4 ) : 1405-1417
- [ 2 ] Wang W N, Liu W Q, Chen H. Information granules-based BP neural network for long-term prediction of time series [ J ]. IEEE Transactions on Fuzzy Systems, 2021, 29 ( 10 ) : 2975-2987
- [ 3 ] Sun P, Aljeri N, Boukerche A. Machine learning-based models for real-time traffic flow prediction in vehicular networks [ J ]. IEEE Network, 2020, 34 ( 3 ) : 178-185
- [ 4 ] 朱苗苗.基于时间序列模型的网络流量预测研究 [ D ].西安:西安工程大学,2017  
ZHU Miaomiao. Research on network traffic prediction based on time series model [ D ]. Xi'an: Xi'an Polytechnic University, 2017
- [ 5 ] Chen L W, Ho Y F. Centimeter-grade metropolitan positioning for lane-level intelligent transportation systems based on the internet of vehicles [ J ]. IEEE Transactions on Industrial Informatics, 2019, 15 ( 3 ) : 1474-1485
- [ 6 ] Li C, Xu P. Application on traffic flow prediction of machine learning in intelligent transportation [ J ]. Neural Computing and Applications, 2021, 33 ( 2 ) : 613-624
- [ 7 ] Chan R K C, Lim J M Y, Parthiban R. A neural network approach for traffic prediction and routing with missing data imputation for intelligent transportation system [ J ]. Expert Systems with Applications, 2021, 171 : 114573
- [ 8 ] Du B W, Peng H, Wang S Z, et al. Deep irregular convolutional residual LSTM for urban traffic passenger flows prediction [ J ]. IEEE Transactions on Intelligent Transportation Systems, 2020, 21 ( 3 ) : 972-985
- [ 9 ] 武佳琪,李珂,檀亚宁.基于 LSTM-BP 神经网络的道路拥堵时间智能预测 [ J ]. 数字技术与应用, 2020, 38 ( 5 ) : 64-65  
WU Jiaqi, LI Ke, TAN Yaning. Intelligent prediction of road congestion time based on LSTM-BP neural network [ J ]. Digital Technology & Application, 2020, 38 ( 5 ) : 64-65
- [ 10 ] 李磊,张青苗,赵军辉,等.基于改进 CNN-LSTM 组合模型的分时段短时交通流预测 [ J ]. 应用科学学报, 2021, 39 ( 2 ) : 185-198  
LI Lei, ZHANG Qingmiao, ZHAO Junhui, et al. Short-term traffic flow prediction method of different periods based on improved CNN-LSTM [ J ]. Journal of Applied Sciences, 2021, 39 ( 2 ) : 185-198
- [ 11 ] Lu S Q, Zhang Q Y, Chen G S, et al. A combined method for short-term traffic flow prediction based on recurrent neural network [ J ]. Alexandria Engineering Journal, 2021, 60 ( 1 ) : 87-94
- [ 12 ] 赵俊峰.网络流量预测的研究 [ D ].天津:天津工业大学,2013  
ZHAO Junfeng. Research on network traffic prediction [ D ]. Tianjin: Tianjin Polytechnic University, 2013
- [ 13 ] 黎佳玥,赵波,李想,等.基于深度学习的网络流量异常预测方法 [ J ]. 计算机工程与应用, 2020, 56 ( 6 ) : 39-50  
LI Jiayue, ZHAO Bo, LI Xiang, et al. Network traffic anomaly prediction method based on deep learning [ J ]. Computer Engineering and Applications, 2020, 56 ( 6 ) : 39-50

- [14] 王海宁,袁祥枫,杨明川.基于 LSTM 与传统神经网络的网络流量预测及应用[J].移动通信,2019,43(8):37-44  
WANG Haining, YUAN Xiangfeng, YANG Mingchuan. Network traffic prediction and application based on LSTM and traditional neural networks[J]. Mobile Communica-
- [15] Liu S Y, Li Z, Li H F. Research on short-term traffic flow prediction model based on RNN-LSTM [J]. IOP Conference Series: Materials Science and Engineering, 2020, 806: 012017

## Prediction of time series traffic based on improved LSTM algorithm

GUO Jiali<sup>1</sup> XING Shuangyun<sup>1</sup> LUAN Hao<sup>1</sup> JIA Yanting<sup>2</sup>

1 School of Sciences, Shenyang Jianzhu University, Shenyang 110168

2 Teaching Affairs Department, Shenyang Jianzhu University, Shenyang 110168

**Abstract** The prediction of time series traffic is a hot issue in machine learning in recent years. It has been found that the prediction accuracy can be greatly improved by approaches of changing the network structure (such as the number of neural network layers, the number of neurons in network layers, the connection mode between network layers, as well as the application of special network layers), and selecting appropriate optimizer and loss function. Here, we propose a multi-layer LSTM (Long Short-Term Memory) algorithm, which is a single model improved on traditional LSTM algorithm, to reduce the model's complexity and improve the efficiency of machine learning. The model includes an input layer, five hidden layers, an output layer, a full connection layer, and also a dropout layer to prevent the machine learning from over-fitting. The model uses adam as optimizer, mse as loss function, and relu as activation function. The experimental results show that the proposed model has better generalization ability compared with traditional LSTM model.

**Key words** time series prediction; long short-term memory (LSTM); prediction accuracy; generalization ability