

基于注意力机制的航班配餐备份数预测算法

王 俊, 吴子轩, 董 杰, 蔡哲立, 张 凯, 许乔若

(厦门航空有限公司, 福建 厦门 361000)

✉ wangjun6@xiamenair.com; zixuan_wu@sina.com; dongjie3@xiamenair.com;
caizheli@xiamenair.com; zhangkai7@xiamenair.com; xuqiaoruo@xiamenair.com



摘 要: 航班配餐备份量受到历史订票人数和多种外部因素的影响, 可以利用注意力机制捕获航班历史订票人数对最终离港人数的影响程度, 通过编解码器结构捕获外部复杂因素与航班旅客订票序列间的关系, 再引入残差设计加快模型的收敛速度并防止梯度爆炸问题。文章提出一种基于注意力机制的航班配餐备份数预测算法, 并在真实航班旅客数据集上进行实验, 以MAPE(平均绝对百分比误差)为评价指标, 该算法较传统的统计学方法、机器学习方法和深度学习方法, 误差分别降低2.19%、0.42%和3.02%。

关键词: 航班配餐备份数; 编解码器; 注意力机制; 残差设计; 深度学习

中图分类号: TP39 **文献标识码:** A

An Algorithm for Predicting the Flight Catering Backups Quantity based on Attention Mechanism

WANG Jun, WU Zixuan, DONG Jie, CAI Zheli, ZHANG Kai, XU Qiaoruo

(XIAMEN AIR, Xiamen 361000, China)

✉ wangjun6@xiamenair.com; zixuan_wu@sina.com; dongjie3@xiamenair.com;
caizheli@xiamenair.com; zhangkai7@xiamenair.com; xuqiaoruo@xiamenair.com

Abstract: The backups quantity of flight catering is affected by the number of historical bookings and various external factors. The attention mechanism can be used to capture the different influence grade of the number of flight history bookings on the final departures. The Encoder-Decoder structure can capture the relationship between the external complex factors and the flight passenger booking sequence. The residual design is introduced to accelerate the convergence speed of the model and prevent the gradient explosion problem. This paper proposes an algorithm for predicting the flight catering backups quantity based on attention mechanism. Taking MAPE (Mean Absolute Percentage Error) as the evaluation index, the empirical studies on real flight passenger datasets show that compared with the traditional statistical methods, machine learning methods and deep learning methods, the proposed algorithm reduces the error by 2.19%, 0.42% and 3.02%, respectively.

Keywords: flight catering backups quantity; encoder-decoder; attention mechanism; residual design; deep learning

1 引言(Introduction)

航班旅客配餐是民航领域的子业务, 是航空产业链的重要环节。配餐备份量受到航班运行情况、旅客订座情况、特殊日期等因素的影响, 业内使用人工经验确定配餐备份量容易导致餐食浪费, 从而增加企业的成本。因此, 准确预测航班配餐备份数有助于航空公司的配餐部门对餐食备份数量的

决策提供支持, 提升服务质量。航班配餐备份数预测, 本质上是一个时间序列预测问题, 一直以来受到国内外学者的广泛研究^[1], 传统数学统计学方法要求数据满足某些假设, 故难以达到预期目标, 传统机器学习方法则需要做极其复杂的特征工程工作。近年, 深度学习方法如循环神经网络(RNN)^[2], 采用编码器-解码器结构的Seq2Seq模型^[3]以及注意力机制^[4]等应

用在时间序列预测问题上取得了很大的成功。本文提出了一种基于注意力机制,并融合编解码器结构和残差设计的模型(ARSeq-Net),用于解决航班配餐备份数的预测问题。

2 研究现状及本文贡献(Related work and contribution)

以民航为研究对象的文献,例如:林友芳等^[6]使用时空长短期记忆网络对出发地—目的地的客运需求进行预测。邓玉婧等^[5]使用基于多粒度时间注意力机制的循环神经网络模型对未来航班的离港人数进行预测。然而,这些研究中针对的具体问题与本文有较大区别,例如林友芳等^[5]预测的是未来出发地—目的地的客运需求,邓玉婧等^[6]专注于用历史起飞时刻的客座率预测航线未来的客座率,而本文则是采用航班起飞前的历史订座数据预测起飞时的客座人数,即配餐备份数。在交通领域的其他时间序列预测问题的研究中,早期主要采用传统数学统计学方法。STEPHANEDES^[7]将历史平均(Historic Average, HA)方法应用于城市交通控制系统。陈杨等^[8]将指数平滑(Exponential Smoothing, ES)法应用于交通流数据预测。HAMED等^[9]使用差分自回归移动平均(AutoRegressive Integrated Moving Average, ARIMA)模型对交通流量进行预测。WILLIAMS等^[10]使用季节性差分自回归移动平均(Seasonal AutoRegressive Integrated Moving Average, SARIMA)模型通过差分将非平稳时间序列转变为平稳时间序列进行交通流量预测。ZIVOT等^[11]考虑空间相关性对交通流量的影响,使用向量自回归(Vector AutoRegression, VAR)模型在交通流量预测上取得了更好的预测精度。

需要注意的是,由于实际中影响航班离港数量的外部条件和因素过于复杂,因此在本文研究的问题中,往往无法满足此类传统数学统计学方法的平稳性假设限定,并且这些方法适用于相对较小的数据集,对于海量的航班数据,无法达到预期的目标。

随着人工智能的发展,传统机器学习方法能够更灵活地刻画交通数据。ALAM等^[12]使用线性回归对交通流量进行预测。姚智胜等^[13]使用支持向量回归方法对交通流量、占有率、平均速度等特征进行训练,用于解决交通数据预测问题。DAVIS等^[14]利用 k 近邻模型解决交通数据的预测问题。HAMNER等^[15]使用随机森林训练移动和停止的车辆数量、路段平均速度等特征对未来交通流量进行预测。这些传统机器学习方法虽然已经能够处理高维数据,并能捕获其复杂的非线性关系,但是仍然需要人工构造复杂的特征工程,以提升模型的性能。

近几年,随着深度学习的快速发展,深度学习被更多地应用于时空交通数据的学习与预测任务中。例如, HUANG等^[16]使用深度信念网络(Deep Belief Networks, DBN)和 YANG等^[17]使用堆叠自编码器(Stacked AutoEncoder, SAE)模型自动学习交通数据的特征,不需要像机器学习模型那样

采用烦琐的特征工程。MA等^[18]和 ZHAO等^[19]使用长短期记忆网络(Long Short-Term Memory, LSTM)预测短期的交通速度和流量。SHI等^[20]针对时空数据预测问题,用卷积操作替换传统LSTM模型中的线性变换操作,使得模型可以同时捕获数据在时间和空间维度上的特征。ZHANG等^[21]提出了ST-ResNet模型,通过卷积层和残差结构的组合,建模了城市范围内的人群流量的空间相关性。YU等^[22]提出时空图卷积网络(Spatio-Temporal Graph Convolutional Network, STGCN)模型,使用一维卷积刻画了时间动态相关性,使用图卷积神经网络(GCN)获取路网的局部空间相关性。

Seq2Seq作为RNN的变种,包括编码器(Encoder)和解码器(Decoder)两部分,在自然语言处理和时间序列任务中都取得了良好的效果。路宽等^[23]使用多层Bi-GRU神经元为基本单位构建Seq2Seq对短期电力负荷进行预测。陶涛等^[24]使用基于双向LSTM的Seq2Seq模型对加油站时序数据进行异常检测。朱墨儒等^[25]使用Seq2Seq模型对于云资源负载进行预测。王涛等^[26]使用Seq2Seq模型并融合外部特征对港口进出口货物量进行预测。RNN采用类似递归的方式捕获序列的信息,无法并行化处理信息,并且存在梯度爆炸和梯度消失的问题。VASWANI等^[4]提出的Transformer模型采用注意力机制解决了该问题,该模型能够更好地捕获时间间隔较长的系列信息。ZHOU等^[27]在Transformer模型的基础上进行了改进,提出了Informer,降低了注意力机制的计算复杂度,提升了在长序列时间预测问题上的预测效果,但此类方法的时间复杂度依然很高。

对于航班配餐备份数预测问题,研究人员希望模型能够自适应学习航班历史订票人数对最终离港人数的影响程度,同时捕获外部复杂因素与航班旅客订票序列间的关系,因此本文提出了一种基于注意力机制和残差设计的编解码器深度模型用于解决这一问题。本文的创新点和具体贡献总结如下。

(1)通过注意力机制能有效捕获航班历史订票人数对最终离港人数的影响程度,从而提高模型的预测精度。此外,本文仅在最后一步隐藏状态以及历史隐藏状态应用了注意力机制,其时间复杂度仅与序列长度呈线性关系,较Transformer和Informer相比,时间复杂度更低。

(2)通过编解码器结构能有效捕获外部复杂因素与航班旅客订票序列间的关系。

(3)在真实数据集上进行了实验,证明了方法的有效性和先进性。

3 问题描述(Problem formulation)

本节对航班配餐备份数预测问题进行形式化描述和定义,并对相关数据特征进行描述。某航班订票人数的变化曲线如图1所示,其中横轴代表时间(STD代表起飞时刻,STD-3h则代表起飞前3 h),纵轴代表航班人数。航班订票人数在时间轴上变化趋势具有连续特性,是典型的时间序列,而航班配餐备份数的预测问题,本质上是预测该航班实际离港人数,

所以本文研究的问题为典型的时间序列预测问题。

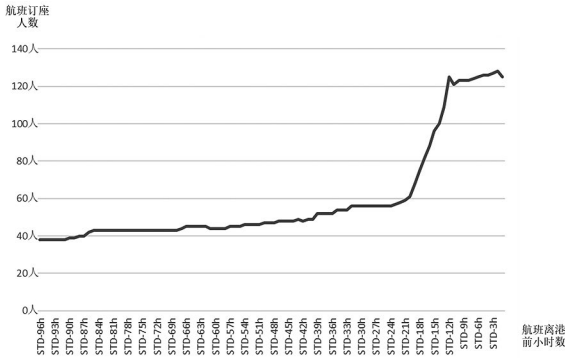


图1 某航班人数变化曲线图

Fig.1 Change curve of the number of people on a certain flight

(1)序列数据输入：本文使用航班起飞前96 h到起飞前2 h的订票人数作为序列输入数据 $\{X_t|t=96,95,\dots,2\}$ 。

(2)外部复杂因素输入：本文使用航班起飞及降落时间、起飞及降落机场、飞机型号、各舱型座位总数，以及航班起飞当日是周几、是否处于法定节假日期间、是否为补班、是否处于春运和暑运期间拼接作为外部复杂因素输入 X_{other} 。

本文的目标：将序列数据和外部复杂因素共同输入模型，以预测离港人数 $F(X, X_{other}) \rightarrow Y$ 。

4 基于注意力机制和残差的深度编解码器网络 (Deep encoder-decoder network based on attention mechanism and residual)

本节将详细介绍基于注意力机制和残差设计的编解码器深度模型(ARSeq-Net)的结构。该模型将注意力机制、残差设计以及编解码器结构应用于航班配餐备份数预测问题，旨在自适应地提取序列中的隐藏信息以及外部因素的影响，通过神经网络的记忆能力、容错能力、自学习能力拟合预测函数，从而进行高效、准确的预测。其中，注意力机制是一种对不同特征信息计算不同权重，并按权重值融合特征信息的方法，能够有效捕获到关键的信息；残差设计能有效解决梯度爆炸和梯度消失问题，加速模型收敛；编解码器结构可以很有效地对输入序列进行特征表示，并捕获到外部因素的影响。

4.1 编解码器

编解码器(Encoder-Decoder)的提出基于Seq2Seq模型，是一种循环神经网络的变种，最早应用于自然语言处理领域。编码器部分的主要作用是任意长度的序列信息编码到传递向量中；解码器的作用主要是将传递向量中所蕴含的上下文信息进行解码，并输出新的序列；其结构图如图2所示。

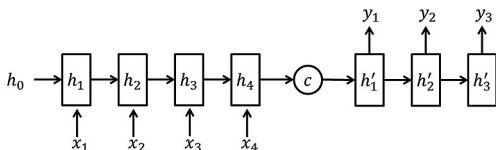


图2 编码器-解码器结构

Fig.2 Encoder-Decoder structure

4.2 注意力机制

注意力机制能够在大量的信息中有效地捕获较为关键的信息，以此提高网络的效率。具体来说，注意力机制就是要对特征信息计算权重，并对所有信息按权重进行合并。注意力机制的计算过程分为两步，第一步是计算历史时间步输出向量表示和最后一步隐藏状态表示的注意力分布，第二步是根据注意力权重的分布情况计算历史时间步向量表示的加权平均。本文中所用到的注意力机制是参考VASWANI等^[4]设计的；具体公式如式(1)所示，具体结构示意图如图3所示。

$$Att(Q, K, V) = softmax(\frac{(QK^T)}{\sqrt{d_k}}V) \quad (1)$$

其中， Q 表示查询， K 表示键值， V 表示输入值， d_k 表示 Q 和 K 的维度。

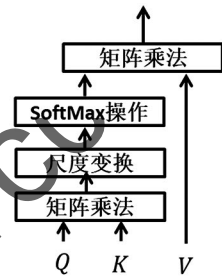


图3 注意力机制示意图

Fig.3 Diagram of attention mechanism

本文中 Q 为编码器最后一步的隐藏状态的线性变换向量， K 和 V 均是编码器每一步输出的线性变换向量。首先通过 Q 和 K 的点乘，即可计算编码器最后一步(即当前步)的隐藏状态与历史时间步输出的相似度。其次通过缩放和SoftMax操作，即可计算出注意力权重矩阵，即序列历史时间步对当前步的影响程度，此时各历史时间步的影响程度数加和为1。最后将注意力权重与 V 进行点乘操作，所得输出即为考虑了不同历史时间步对当前的不同影响后的融合信息表示。这样就能通过模型自适应地学习不同历史时间对于当前隐藏状态的影响，提升模型的预测效果。

4.3 残差设计

残差网络(Residual Network, ResNet)是为了解决卷积神经网络隐藏层过多时的网络退化问题而提出的，但是它并不局限于卷积神经网络，其通过给非线性变换层增加直连边提高信息的传播效率，防止梯度爆炸和梯度消失，以及加快模型收敛速度。

当输入为 x 时，假设神经网络学习到的特征为 $H(x)$ ，此时通过残差单元学习的内容可记作 $F(x) = H(x) - x$ ，可以使网络更容易学习。

本文将残差设计引入编解码器结构，旨在防止梯度爆炸，加快模型的收敛速度，提升模型预测的精度。

4.4 编解码器深度航班配餐备份数预测模型

本文提出了一种基于注意力机制和残差设计的编解码器

深度模型(ARSeq-Net),用于航班配餐备份数的预测,其设计架构如图4所示,包括编码器、注意力机制模块、解码器以及全连接模块。其中 x 表示模型输入, Y' 表示模型输出, i 表示样本数。图4上半部分左右两个虚线框为编码器(Encoder)和解码器(Decoder),两者均采用单层循环门控单元(GRU)进行构建,其中虚线箭头部分表示残差设计。编码器的输入为历史起飞前96 h到起飞前2 h的航班旅客订票统计数。图4下半部分为注意力机制模块,该模块输入为编码器的输出 o^i 以及最后一步隐藏参数 h_{t-1}^i ,输出为 $o_{t-1}^{i'}$ 。经过编码器得到的隐藏层表示 h_{t-1}^i 与不同时间步的输出 o^i 经过注意力机制的计算后与其他特征 x_{other}^i 进行拼接后,经过解码器单元进行解码得到最终的预测值 Y' 。

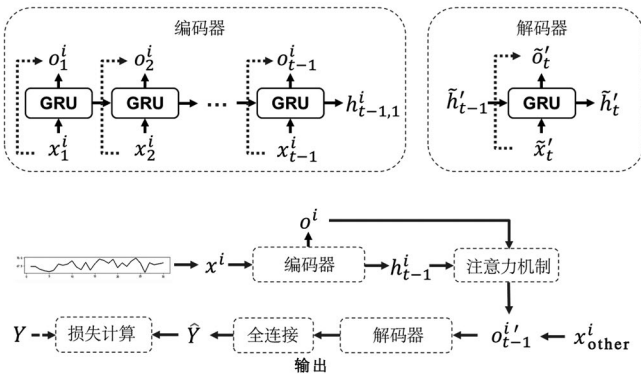


图4 ARSeq-Net算法模型架构图

Fig.4 ARSeq-Net algorithm structure

相比其他模型,ARSeq-Net在捕获历史依赖关系的基础上,通过注意力机制,能够自适应地学习历史对当前的不同影响程度,通过残差设计,有效提升了模型的收敛速度,并很好地解决了梯度爆炸等问题,使得预测结果更准确。

5 实验结果与分析(Experimental results and analysis)

5.1 数据集

本文实验使用的数据集来自某航空公司提供的真实航班旅客订座数据集,数据集中共计涉及的机场数量为105个,其中同时作为起飞机场和降落机场的84个,仅作为起飞机场的10个,仅作为降落机场的11个,共计存在的机场对为377个,一共有5,021个不同的航班;时间跨度从2018年3月16日到2020年12月31日,共计11,109,789条订座数据。

5.2 数据预处理

(1)划分训练集将原始数据处理为序列数据起飞前96 h到起飞前2 h的订座人数为序列特征、起飞后的订座人数为标签值,共得到序列数据432,029条,按每条航班的7:2:1的比例划分训练集、验证集、测试集,其中训练集包含85,786条数据,验证集包含24,739条数据,测试集包含13,482条数据。

(2)数据标准化。对航班订座人数进行z-score标准化,通过标准化可以加快模型收敛速度,提升模型精度。

5.3 基准方法与评价指标

经过调研,目前与本文研究解决同一类问题的方法极少。基于此,本文选定了时间序列预测的经典方法以及近年来先进的代表性方法,共计12个基准方法与ARSeq-Net进行了实验和比较。

HA:历史平均值模型(Historical Average),是一种使用历史平均值对未来进行预测的统计学方法。

ES:指数平滑预测法模型(Exponential Smoothing),是一种引入计算权重,并将计算权重由近到远按指数规律递减的一种特殊的加权平均的统计学方法。

LR:线性回归模型(Linear Regression),是一种利用线性回归方程的最小平方函数对多个自变量和因变量之间关系进行建模和预测的机器学习方法。

KNR:K近邻回归模型(K-Neighbors Regressor),是一种利用K个最近训练样本的目标数值,对待测样本的回归值进行决策的机器学习方法。

DTR:决策树回归模型(Decision Tree Regressor),是一种以信息熵为度量构造下降最快的树的自顶向下的递归的机器学习方法。

SVR:支持向量回归模型(Support Vector Regression),是一种运用支持向量机(SVM)拟合曲线并进行回归分析的机器学习方法。

RFR:随机森林回归模型(Random Forest Regressor),是一种利用生成多决策树的方式对样本进行训练并预测的机器学习方法。

XGBoost:是一种利用多个梯度提升决策树(GBDT)进行训练并计算每个子树的权重进行累加的机器学习集成模型。

LSTM:长短期记忆模型,是一种更适合处理长时间序列数据的循环神经网络模型。

GRU:门控循环单元,是一种参数量较少,更容易收敛的循环神经网络模型。

Transformer:是一种基于多头注意力机制的编码解码长序列预测模型。

Informer:是一种基于Transformer改进得更为高效的长序列预测模型,专门适用于时间序列预测问题。

本文采用平均绝对误差(MAE)、均方根误差(RMSE)和平均绝对百分比误差(MAPE)作为评价指标,具体计算如式(2)至式(4)所示:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (3)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \times 100\% \quad (4)$$

其中, $\hat{y}_i$ 和 y_i 分别表示预测值和真实值, n 表示预测样本总数。

5.4 实验结果

在真实航班旅客数据集上, 验证了ARSeq-Net模型的预测效果。模型的超参设置如表1所示, 预测结果如表2所示。

表1 ARSeq-Net的超参设置

Tab.1 Superparametric setting of ARSeq-Net

参数	值	备注
Hidden_dim	128	隐藏层神经元个数
Depth	1	网络层数
SequenceLength	95	输入序列长度
ResultLength	1	返回序列长度
Normalise	true	正则化
Optimizer	Adam	优化器
Batchsize	128	批大小

表2 不同方法在数据集上实验结果比较

Tab.2 Comparison of experimental results of different methods on data sets

模型	MAE	RMSE	MAPE/%
HA	23.12617	31.96331	30.17332
ES	6.43467	11.49825	6.26116
LR	6.25965	9.29753	7.26443
KNR	8.00021	11.13612	9.29720
DTR	6.33014	9.99240	6.78505
SVR	6.63477	10.26274	8.07075
RFR	4.38336	7.12092	5.04199
XGBoost	4.11794	6.69165	4.49381
LSTM	6.06474	9.79363	7.09497
GRU	6.08005	10.97935	8.26730
Transformer	5.68696	9.87693	7.24921
Informer	5.77253	10.02069	7.30183
ARSeq-Net	3.99816	6.61584	4.0706

表2显示了不同模型在数据集上的预测结果, 本文的ARSeq-Net模型的MAE、RMSE和MAPE都优于其他模型。以MAPE为例, ARSeq-Net较传统的统计学方法中最佳的ES, 误差降低了2.19%; 较机器学习方法中最佳的XGBoost, 误差降低了0.42%; 较深度学习方法中最佳的LSTM, 误差降低了3.02%。其中, HA和ES的效果较差, 这是由于统计学方法并不能很好地建模序列预测问题; LR、KNR、DTR、SVR等模型的效果有所提升, 但是它们不能很好地捕获时间序列的特征; LSTM和GRU作为循环神经网络模型, 虽然效果有所提升, 但是对于长时间序列建模的效果不是很好; 而Transformer和Informer虽然采用了注意力机制, 但是此类方法在编码器中采用自注意力机制进行计算, 再通过注意力机制与解码器进行连接, 需要对两两时间步进行注意力权重计算, 时间复杂度较高, 信息也容易出现的过度冗余, 导致预

测效果不佳。

5.5 消融实验

为了进一步验证模型中每个组件的有效性, 研究人员对ARSeq-Net的三种变体进行了实验。

(1)GRU: 不使用编解码器结构的模型。

(2)Seq2Seq: 仅使用编解码器结构的模型。

(3)ASeq-Net: 使用了编解码器结构和注意力机制的模型。

实验结果如表3所示。

表3 消融实验结果对比

Tab.3 Comparison of ablation experiment results

模型	MAE	RMSE	MAPE/%
GRU	6.08005	10.97935	8.26730
Seq2Seq	4.37094	7.16131	5.90921
ASeq-Net	4.32496	7.10278	4.65113
ARSeq-Net	3.99816	6.61584	4.0706

从表3的实验结果可以看出, ARSeq-Net模型的三个评价指标都是最好的, GRU模型的效果最差。实验结果可以证明, 在加入了编解码器结构后, 模型效果有了一定提升; 加入注意力机制后, 效果又有了进一步提升; 残差的进一步引入, 使得模型效果达到最好, 由此可以证明ARSeq-Net模型的有效性。

6 结论(Conclusion)

本文提出了一种基于注意力机制和残差设计的Seq2Seq深度网络模型, 用于解决航班配餐备份数预测问题, 研究人员将本文提出的模型同处理时间序列预测的传统数学统计学方法、经典的机器学习方法以及经典和先进的深度学习方法在相同数据集上进行了实验, 实验结果表明, 本文提出的模型在预测效果上达到了最优, 证明了本方法的有效性。但是, 该方法还存在一些优化空间, 比如可以考虑其他外部因素如疫情等特殊时期的影响、天气、经济形势变化等对航班配餐备份数的影响, 此部分因素难以合理量化, 可作为未来模型优化的方向。

参考文献(References)

[1] 刘万明.对航空配餐问题的量化分析[J].民航学报,2020,4(05): 1-4.

[2] ELMAN J L. Finding Structure in Time[J]. Cognitive Science, 1990, 14(2):179-211.

[3] SUTSKEVER I, VINYALS O, LE Q V. Sequence to sequence learning with neural networks[C]// ZOUBIN G, MAX W, CORINNA C, et al. Advances in Neural Information Processing Systems. Montreal: MIT Press, 2014:3104-3112.

[4] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]// ISABELLE G, ULRIKE V, SAMY B, et al. Advances in Neural Information Processing Systems. Long Beach: MIT Press, 2017:6000-6010.

- [5] 林友芳,尹康,党毅,等.基于时空LSTM的OD客运需求预测[J]. 北京交通大学学报,2019,43(1):114-121.
- [6] 邓玉婧,武志昊,林友芳.基于多粒度时间注意力RNN的航班客座率预测[J]. 计算机工程,2020,46(01):294-301.
- [7] STEPHANEDES Y J. Introduction: Special Section on Real-Time Traffic Control System[J]. Journal of Transportation Engineering, 1990, 116(6):701-702.
- [8] 陈杨,刘光勇,唐海周,等.一种基于平衡指数平滑的交通流数据预测方法[J]. 现代计算机,2014,20(7):45-48.
- [9] HAMED M M, AL-MASAEID H R, SAID Z M B. Short-term prediction of traffic volume in urban arterials[J]. Journal of Transportation Engineering, 1995, 121(3):249-254.
- [10] WILLIAMS B M, HOEL L A. Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results[J]. Journal of Transportation Engineering, 2003, 129(6):664-672.
- [11] ZIVOT E, WANG J. Vector autoregressive models for multivariate time series[M]. Berlin: Springer, 2003:385-429.
- [12] ALAM I, FARID D M, ROSSETTI R J F. The prediction of traffic flow with regression analysis[M]. Emerging Technologies in Data Mining and Information Security. Singapore: Springer, 2019:661-671.
- [13] 姚智胜,邵春福,高永亮.基于支持向量回归机的交通状态短时预测方法研究[J]. 北京交通大学学报,2006,30(3):19-22.
- [14] DAVIS G A, NIHAN N L. Nonparametric regression and short-term freeway traffic forecasting[J]. Journal of Transportation Engineering, 1991, 117(2):178-188.
- [15] HAMNER B. Predicting travel times with context-dependent random forests by modeling local and aggregate traffic flow[C]// WEI F, WYNNE H, GEOFFREY I, et al. Proceedings of the 2010 IEEE International Conference on Data Mining Workshops. Washington D.C.: IEEE Computer Society, 2010:1357-1359.
- [16] HUANG W, SONG G, HONG H, et al. Deep architecture for traffic flow prediction: deep belief networks with multitask learning[J]. IEEE Transactions on Intelligent Transportation Systems, 2014,15(5):2191-2201.
- [17] YANG H F, DILLON T S, CHEN Y P P. Optimized structure of the traffic flow forecasting model with a deep learning approach[J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(10):2371-2381.
- [18] MA X, TAO Z, WANG Y, et al. Long short-term memory neural network for traffic speed prediction using remote microwave sensor data[J]. Transportation Research Part C: Emerging Technologies, 2015, 54(5):187-197.
- [19] ZHAO Z, CHEN W, WU X, et al. LSTM network: a deep learning approach for short-term traffic forecast[J]. IET Intelligent Transport Systems, 2017, 11(2):68-75.
- [20] SHI X, CHEN Z, WANG H, et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting[C]// CORINNA C, NEIL D, LEE D, et al. Proceedings of Neural Information Processing Systems. New York: Curran Associates Inc, 2015:802-810.
- [21] ZHANG J, ZHENG Y, QI D. Deep spatio-temporal residual networks for citywide crowd flows prediction[C]// SATINDER S, SHAUL M. Proceedings of AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2017: 1655-1661.
- [22] YU B, YIN H, ZHU Z. Spatio-Temporal Graph Convolutional Networks: A Deep Learning Framework for Traffic Forecasting[C]// LANG J. Proceedings of International Joint Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2018:3634-3640.
- [23] 路宽,孟祥荣,孙雯雪,等.多层Bi-GRU的Seq2seq网络短期电力负荷预测模型[C]//中国光学工程学会.2017智能电网新技术发展与应用研讨会论文集.北京:中国电力科学研究院,2017:138-144.
- [24] 陶涛,周喜,马博,等.基于双向LSTM的Seq2Seq模型在加油站时序数据异常检测中的应用[J]. 计算机应用,2019,39(003): 924-929.
- [25] 朱墨儒,高仲合.用于云资源负载预测的Seq2seq模型[J]. 通信技术,2020,53(1):115-119.
- [26] 王涛,张伟,贾宇欣,等.基于Seq2Seq模型的港口进出口货物量预测[J]. 计算机系统应用,2020,29(3):136-143.
- [27] ZHOU H, ZHANG S, PENG J, et al. Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting[C]// NGUYEN T, HSU D, LEE W, et al. Proceedings of the AAAI Conference on Artificial Intelligence. Palo Alto: AAAI Press, 2021, 35(12):11106-11115.

作者简介:

王 俊(1988-),男,硕士,工程师.研究领域:信息技术,企业数字化,软件测试.

吴子轩(1991-),男,博士,工程师.研究领域:数据科学与人工智能,企业数字化.本文通信作者.

董 杰(1996-),男,本科,工程师.研究领域:软件开发,产品管理.

蔡哲立(1989-),男,本科,工程师.研究领域:软件开发,信息技术,项目管理.

张 凯(1994-),男,本科,工程师.研究领域:软件开发,信息技术.

许乔若(1992-),男,硕士,高级工程师.研究领域:信息技术,数据开发,数据管理.