



计算机科学

COMPUTER SCIENCE

基于LSTM和注意力机制的远程会诊需求预测

翟运开, 乔正文, 乔岩

引用本文

翟运开, 乔正文, 乔岩. 基于LSTM和注意力机制的远程会诊需求预测[J]. 计算机科学, 2024, 51(6A): 230800119-7.

ZHAI Yunkai, QIAO Zhengwen, QIAO Yan. [Forecasting Teleconsultation Demand Based on LSTM and Attention Mechanism](#) [J]. Computer Science, 2024, 51(6A): 230800119-7.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于SAMNV3的滚动轴承智能故障诊断方法](#)

Intelligent Fault Diagnosis Method for Rolling Bearing Based on SAMNV3

计算机科学, 2024, 51(6A): 230700167-6. <https://doi.org/10.11896/jsjcx.230700167>

[融入类别标签和主题信息的用户兴趣识别方法](#)

User Interest Recognition Method Incorporating Category Labels and Topic Information

计算机科学, 2024, 51(6A): 230500169-8. <https://doi.org/10.11896/jsjcx.230500169>

[基于注意力的多尺度蒸馏异常检测](#)

Attention-based Multi-scale Distillation Anomaly Detection

计算机科学, 2024, 51(6A): 230300223-11. <https://doi.org/10.11896/jsjcx.230300223>

[基于改进Deeplabv3+算法的滚珠丝杠驱动表面点蚀缺陷检测](#)

Detection of Pitting Defects on the Surface of Ball Screw Drive Based on Improved Deeplabv3+ Algorithm

计算机科学, 2024, 51(6A): 240200058-6. <https://doi.org/10.11896/jsjcx.240200058>

[融合注意力机制的DeeplabV3+服装图像分割方法](#)

Clothing Image Segmentation Method Based on Deeplabv3+ Fused with Attention Mechanism

计算机科学, 2024, 51(6A): 230900153-7. <https://doi.org/10.11896/jsjcx.230900153>

基于 LSTM 和注意力机制的远程会诊需求预测

翟运开^{1,2,3} 乔正文¹ 乔岩¹

1 郑州大学管理学院 郑州 450001

2 互联网医疗系统与应用国家工程实验室 郑州 450052

3 河南省智能健康信息系统国际联合实验室 郑州 450000

(zhaiyunkai@zzu.edu.cn)

摘要 为更准确地预测远程会诊需求量,提高远程会诊资源配置效率,文中引入多元回归分析(Multiple Linear Regression)和注意力机制来优化长短期记忆网络(LSTM)。首先,根据远程会诊需求中存在的假期效应生成假期指标,通过多元回归分析选取显著性高的指标作为模型输入,然后根据长短期记忆网络学习输入指标的内部复杂映射关系,利用注意力机制对指标分配不同权重,最后根据权重和 LSTM 隐藏层输入预测结果。基于国家远程医疗中心(NTCC)的实际历史会诊数据,研究 MLR-Attention-LSTM 的预测性能,并比较其与整合移动平均自回归模型、支持向量机、K 近邻、BP 神经网络和 LSTM 神经网络 5 种模型的预测效果。结果表明,优化后的 LSTM 模型预测精度最高。进一步地,探究假期指标对模型性能的影响,结果表明假期指标的输入可以进一步提高模型的预测精度,验证了 MLR-Attention-LSTM 和假期相关变量输入在远程会诊需求预测领域的可行性与适用性,为远程医学中心实际应用提供了理论支撑和实践指导。

关键词 长短期记忆网络;注意力机制;远程会诊;需求预测;假期效应

中图分类号 TP183

Forecasting Teleconsultation Demand Based on LSTM and Attention Mechanism

ZHAI Yunkai^{1,2,3}, QIAO Zhengwen¹ and QIAO Yan¹

1 School of Management, Zhengzhou University, Zhengzhou 450001, China

2 National Engineering Laboratory of Internet Medical Systems and Applications, Zhengzhou 450052, China

3 Henan Province International Joint Laboratory of Intelligent Health Information System, Zhengzhou 450000, China

Abstract To predict the demand for teleconsultation more accurately and improve the efficiency of resource allocation for teleconsultation, this paper introduces multiple linear regression and attention mechanism to optimize Long Short-term Memory network. Firstly, according to the holiday effect existing in the teleconsultation demand, the holiday index is generated, and the index with high significance is selected as the model input through multiple regression analysis. Then, according to the long-term short-term memory network to learn the internal complex mapping relationship of the input indicators, the attention mechanism is used to assign different weights to the indicators. Finally, the prediction results are input according to the weight and LSTM hidden layer. Based on the actual historical teleconsultation data of the National Telemedicine Center, this paper studies the predictive ability of MLR-Attention-LSTM, and compares it with the ARIMA, SVR, KNN, BP neural network and long short-term memory network. The results show that the improved LSTM model has the highest prediction accuracy. Furthermore, this paper explores the impact of holiday indicators on the performance of the model. The results show that the input of holiday indicators can further improve the prediction accuracy of the model. It verifies the feasibility and applicability of MLR-Attention-LSTM and holiday-related variable input in the field of teleconsultation demand prediction, and provides theoretical support and practical guidance for the practical application of telemedicine centers.

Keywords Long short-term memory, Attention mechanism, Teleconsultation, Demand forecasting, Holiday effect

1 引言

党的二十大报告提出,要促进优质医疗资源扩容和区域均衡布局。目前我国医疗资源大多集中于城市,分布不均衡,由此会导致供需不平衡,降低医疗资源使用效率。远程医疗

能有效缓解医疗分布不均的问题,近些年来,远程医疗模式受到了越来越多的关注,政府先后出台多个政策意见引导鼓励医疗机构开展远程医疗。2020年2月,《关于在疫情防控中做好互联网诊疗咨询服务工作的通知》提出要充分发挥互联网医疗服务优势,大力开展互联网诊疗服务。2022年3月3

基金项目:国家自然科学基金(72202217,71972012);河南省高等学校重点科研项目计划(24A630034);河南省高等学校哲学社会科学基础研究重大项目(2022-JCZD-21)

This work was supported by the National Natural Science Foundation of China(72202217,71972012), Key Research Project Plan for Colleges and Universities of Henan Province(24A630034), Major Project of Basic Research of Philosophy and Social Science in Colleges and Universities of Henan Province(2022-JCZD-21).

通信作者:乔岩(zzuqiaoyan@zzu.edu.cn)

日,在国务院关于印发《“十四五”中医药发展规划》的通知中指出,要建设中医互联网医院,发展远程医疗和互联网诊疗。远程会诊是远程医疗服务中的主要业务,是一种基于两个或多个地理上分开的医疗机构之间进行基于医疗数据跨机构共享的远程会商、讨论并给予申请方诊断指导的医疗模式,主要用于诊断或治疗疑难疾病和突发疾病。通过远程会诊,将优质的医疗资源下沉到医疗不发达的基层地区,可以有效缓解我国医疗资源分布不均的问题。

在国家的政策引导下,越来越多的基层医院与国家远程医疗中心联网,远程会诊的需求逐渐增加。精准的远程会诊需求预测能够提高远程会诊的服务效率和运作效率,为远程会诊运营管理水平和服务质量提升提供理论依据,其研究结果具有重要理论价值。同时,本文基于国家远程医疗中心的实际远程会诊数据对远程会诊需求进行预测,研究结果能够直接或间接地优化远程会诊资源的配置和使用,为分诊员的配置提供决策依据,减少供需不均的现象,为后期的决策优化问题提供实践指导,具有重要的实际应用价值。

然而目前很少有文献研究远程会诊的需求预测问题。大多数远程医疗运营方面的研究主要集中在资源配置^[1]、供需匹配^[2]、预约调度^[3]、专家推荐^[4]等方面。在许多领域,如供应链^[5]、交通^[6]、能源^[7]、旅游业^[8],需求预测的应用提高了其运营管理水平和服务质量。近年来,远程会诊需求量逐渐增加,亟需远程会诊需求预测相关研究为远程会诊资源配置提供指导和借鉴。

在医疗健康领域,就诊需求具有假期效应^[9]。假期效应是指在假期前后,医疗机构的各类诊疗服务量会减少,进而医务人员、病房资源、ICU资源等会进行适当调整,降低供给数量,在满足就诊需求的前提下降低机构运行成本。在可视化的远程会诊历史数据中可以观察到,远程会诊也具有明显的假期效应。基于远程会诊主要解决基层复杂与疑难疾病诊断的定位,一方面,在假期中及假期前后,基层医院的诊疗量会出现下降,进而使得远程会诊需求量下降;另一方面,由于基层医院对上级医院放假及其值班专家减少的担忧,在假期前后会集中申请疑难疾病远程会诊,而在假期中仅申请特别紧急的疾病会诊,两种情形使得远程会诊需求预测更加具有挑战性。为了处理存在假期效应下的需求预测,以往研究大致分为两类方法:1)构建高性能模型;2)输入变量处理。

在构建高性能模型方面,传统计量经济学模型对变量间的非线性特征拟合效果较差,而机器学习模型尤其是深度学习对非线性、非平稳的数据拟合效果较好,对变量间非线性关系捕捉能力强,其中 LSTM 神经网络是深度学习中最成熟最具代表性的网络模型。在时间序列预测方面,Ouyang 等^[10]将小波分析与纳入时间序列依赖特征的长短期记忆神经网络相结合,构建金融时间序列数据预测模型,结果表明 LSTM 神经网络具有更高的预测精度。Chen 等^[11]构建基于集成注意力的双向长短期记忆网络,引入不同的输入验证了变量选择的有效性,通过与 9 个基准模型对比,结果表明 LSTM 模型可以得到更好的预测结果。Tang 等^[12]构建 EEMD-LSTM 模型进行保险业系统性风险预警,实证结果表明,EEMD 分解集成技术在预测非线性、非平稳复杂时间序列方面具有明显优势。

在输入变量处理方面,为了从输入变量角度处理假期效应,Liu 等^[13]使用 RF 和 CA-SFS 对 19 个指标进行特征提取,

最后使用 LSTM 模型对股票价格涨跌进行预测,结果表明 LSTM 比浅层机器学习模型预测准确率提高了 33.17%,相比结合 PCA, LASSO 等降维方法的 LSTM 模型准确率提高了 11.45%。Liu 等^[16]将最小二乘支持向量机(LSSVM)和改进的粒子群优化(PSO)算法结合来处理假日客流流的复杂波动从而对假日客流进行预测。Yao 等^[17]引入土地匹配度、风景名胜假期、服务水平等解释性变量,以处理度假旅客出行行为的特殊性,进而预测不同节假日情况下地铁网年变化的客流分布。

综上所述,目前远程会诊相关研究存在如下问题:

1)远程会诊需求预测的研究较少。伴随着远程会诊服务体系纳入我国健康服务体系建设规划,90%以上的县级以上医疗机构已经加入各类远程医疗网络,远程会诊量快速增加,远程会诊平台、优质医疗机构远程会诊服务量倍增,已成为医疗机构的重要服务领域,但也出现了会诊安排不及时、专家资源不足、诊室不足等问题。因此,开展远程会诊需求预测对基层申请医院和优质的专家医院同样重要。

2)融合机器学习、深度学习技术开展远程会诊需求预测的研究尚为空白。远程会诊解决疑难重症和突发急症的定位,开展日级甚至更短时间精度的需求预测更具有理论和实际意义,亟需探讨不同预测模型的实用性和效能、输入变量处理、更短时间精度预测的模型效能,而当前在这方面的研究极为少见。

针对目前存在的问题,本文开展如下工作:

1)构建 LSTM 模型来预测远程会诊的每日需求量,同时基于国家远程会诊中心的历史数据进行实证研究,并使用 4 种模型:差分整合移动平均自回归模型(Autoregressive Integrated Moving Average model, ARIMA)、机器学习模型支持向量机回归(Support Vector Regression, SVR)、最近邻法(K-Nearest Neighbor, KNN)和深度学习模型反向传播神经网络(Back Propagation Neural Network, BPNN)作对比以探究 LSTM 神经网络在远程会诊需求预测中的可行性与适用性。

2)为了进一步提高模型预测准确性,在输入变量的处理方面,从大数据的角度选择假期相关变量添加到输入变量中,从而将更多的有用信息导入到预测模型中。

本文的主要贡献如下:1)将 LSTM 神经网络引入远程会诊需求预测,并建立预测模型,对远程会诊进行“日级”需求预测,以提高预测精度,从而提高远程会诊资源配置效率;2)对远程会诊的假期效应进行更为详细的处理,选择附加变量以提高每日会诊的预测准确性;3)对不同模型进行比较,讨论机器学习和深度学习模型对远程会诊需求预测的适用性,对比评价多种不同模型之间的预测精度和效果。

2 模型构建

2.1 长短期记忆神经网络

长短期记忆神经网络(Long Short-Term Memory, LSTM)是一种特殊的循环神经网络,最早由 Hochreiter 和 Schmidhuber 于 1997 年提出^[18],它可以有效地解决传统 RNN 面临的梯度消失和梯度爆炸问题,从而更好地捕捉时间序列数据中的长期依赖关系。

LSTM 包含了一个记忆单元(Memory Cell),用于存储和更新历史时刻的状态信息,以及 3 个门控单元,输入门、遗忘门和输出门,用于控制信息的流动和遗忘。结构如图 1 所示。

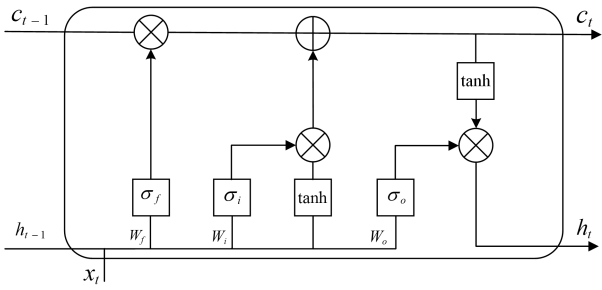


图1 LSTM模型结构

Fig.1 LSTM model structure

由图1知, c_{t-1} 为上一时刻细胞状态, h_{t-1} 为上一时刻隐藏层的值, x_t 为当前 t 时刻输入。

遗忘门: 决定应丢弃或保留哪些信息。来自前一个隐藏状态的信息和当前输入的信息同时传递到 sigmoid 函数中去, 输出值介于 $0 \sim 1$ 之间, 越接近 0 意味着越应该丢弃, 越接近 1 意味着越应该保留。

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (1)$$

其中, σ 为 sigmoid 函数, W_f 和 b_f 分别是遗忘的权重和偏置项, f_t 为遗忘门输出。

输入门: 输入门用于更新细胞状态。

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (2)$$

$$\bar{c}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (3)$$

其中, W_i 和 W_c 为输入门权重, b_i 和 b_c 为输入门的偏置项, $\tanh()$ 为激活函数, i_t 为输入门输出, \bar{c}_t 为候选细胞状态。

细胞状态: 前一层的细胞状态与遗忘向量逐点相乘。

$$c_t = f_t \odot c_{t-1} + i_t \odot \bar{c}_t \quad (4)$$

其中, c_t 为当前时刻细胞状态, \odot 为点乘。

输出门: 输出门用来确定下一个隐藏状态的值, 隐藏状态包含了先前输入的信息。

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t \odot \tanh(c_t) \quad (6)$$

其中, h_t 为 t 时刻输出, W_o 和 b_o 为输出门权重和偏置项。

2.2 注意力机制

注意力机制是指大脑在处理信息时, 选择性地关注某些信息而忽略其他信息的能力。在神经网络中, 注意力机制通常用于提高模型对于输入数据中某些“重要性”高的数据的关注度即权重, 以便更加准确地进行预测或生成输出。具体来说, 通过对输入数据进行加权处理, 注意力机制可以使模型更加关注对当前输出“重要性”高的信息, 从而提高模型的性能。注意力机制最初应用于自然语言处理, 现在已广泛应用于各种时间序列处理。

注意力机制包括多种类型和变体, 其中自注意力机制是注意力机制的一种, 其结构如图2所示。

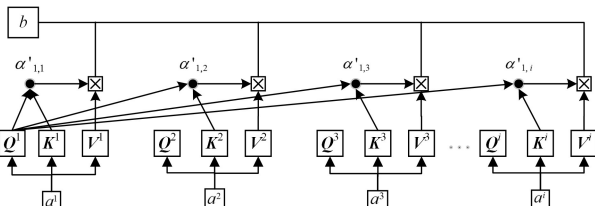


图2 注意力机制结构

Fig.2 Structure of attention mechanism

自注意力机制更关注输入特征内部之间的相关性, 图中

a^1, a^2, a^3, a^i 为输入的特征值; Q, K, V 分别为 Query 值、Key 值和 Value 值。不同于一般的注意力机制, 其 Q, K, V 同源都来自于同一个输入 a , 自注意力机制计算公式如下:

$$b = \sum_{i=1} \text{softmax} \left(\frac{QK^T}{\sqrt{d_k}} \right) V^i \quad (7)$$

其中:

$$Q^i = W^q a^i \quad (8)$$

$$K^i = W^k a^i \quad (9)$$

$$V^i = W^v a^i \quad (10)$$

其中, W^q, W^k, W^v 为可学习的参数, d_k 为输入特征维度。

2.3 Self-Attention-LSTM 神经网络模型

Self-Attention-LSTM 神经网络模型结构如图3所示, 具体结构如下:

1) 输入层。将国家远程医疗中心的历史会诊数据根据远程会诊存在的假期效应选择相关假期, 使用多元回归分析进行指标筛选, 再将数据通过输入层输入到模型中。设输入数据长度为 i , 输入向量 $\mathbf{X} = [x_1, x_2, x_3, \dots, x_i]^T$ 。

2) LSTM 层。经过多元回归分析筛选指标后的数据输入至 LSTM 层。经过内部激活函数进行非线性变换从而学习不同指标间的关系。本文采用二层 LSTM 结构。设输出向量为 \mathbf{a} , 长度为 j , 则 $\mathbf{a} = [a^1, a^2, a^3, a^4, a^i]^T$ 。

3) Attention 层。Attention 层的输入为 LSTM 层的输出 \mathbf{a} , 对 \mathbf{a} 分别如式(8)~式(10)计算 Q, K, V 。通过式(7)计算并求和得到 Attention 层的输出 \mathbf{b} 。

4) 输出层。将 Attention 层输出向量 \mathbf{b} 作为输出层的输入向量, 输出层通过全连接层得到预测值。本文输出值长度为 1, 设输出向量为 \mathbf{y} , 则 $\mathbf{y} = [y^1]^T$ 。

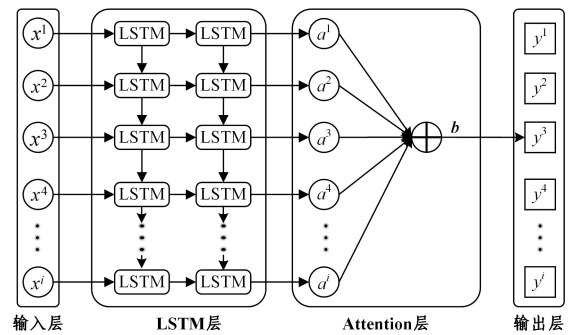


图3 Self-Attention-LSTM神经网络模型结构

Fig.3 Self-Attention-LSTM neural network model structure

2.4 模型评价指标

本文采用均方根误差 RMSE(Root Mean Square Error)、平均绝对误差 MAE(Mean Absolute Error)和 R-squared 这 3 个评价指标来评估模型的预测性能, 公式分别如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (13)$$

RMSE 和 MAE 值越小表明预测误差越小, 模型预测效果越好; R-squared 取值范围为 $[0, 1]$, 其值越接近 1 表明模型对数据的拟合程度越好。

3 实证分析

为研究 MLR-Attention-LSTM 神经网络模型对远程会诊需求量的预测效果,本部分使用 MLR-Attention-LSTM 神经网络模型预测远程会诊需求量,该数据来源于国家远程医疗中心 (NTCC) 数据库的多年远程会诊记录。具体流程图如图 4 所示。

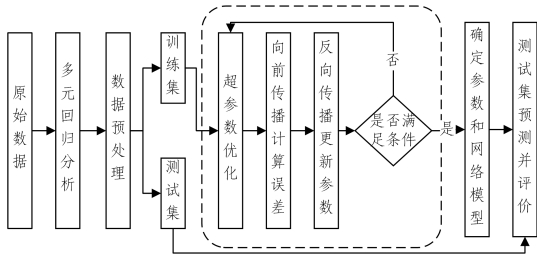


图 4 整体流程图

Fig. 4 Overall flowchart

3.1 数据预处理

3.1.1 数据集介绍

数据集包括 2016 年、2018—2020 年、2021 年 1—6 月的会诊记录,共 1643 条。本研究将数据分为训练集和测试集。对于测试集的选择,为了测试模型对于假期效应的学习效果,应尽可能地包含假期但同时要减少预测时间的跨度,在实际应用中,时间跨度过大会使模型预测结果可信度降低,因此选择 2021 年 1—6 月数据集 (181 天) 作为测试集,数据划分如表 1 所列。

表 1 数据集划分

Table 1 Dataset partition

数据	条数	时间
训练集	1462	2016 年、2018 年—2020 年
测试集	181	2021-01-01—2021-06-30

3.1.2 多元回归分析筛选指标

国家远程医疗中心是国家卫健委批准的依托郑州大学第

表 2 假期指标相关性分析

Table 2 Correlation analysis of holiday indicators

指标	取值	回归系数	显著性	指标	取值	回归系数	显著性
星期	1-7	-12.56	<0.01	星期	1-7	-12.55	<0.01
是否节假日	0.1	-45.93	<0.01	是否节假日	0.1	-45.88	<0.01
节前第三天	0.1	-2.82	0.427	节前第二天	0.1	-12.80	<0.01
节前第二天	0.1	-12.87	<0.01	节前第一天	0.1	-29.78	<0.01
节前第一天	0.1	-29.84	<0.01	节后第一天	0.1	13.67	<0.01
节后第一天	0.1	13.61	<0.01	节后第二天	0.1	4.44	0.214
节后第二天	0.1	4.37	0.220	节后第三天	0.1	6.06	0.089
节后第三天	0.1	6.00	0.093				
指标	取值	回归系数	显著性	指标	取值	回归系数	显著性
星期	1-7	-12.56	<0.01	星期	1-7	-12.57	<0.01
是否节假日	0.1	-46.03	<0.01	是否节假日	0.1	-46.21	<0.01
节前第二天	0.1	-12.89	<0.01	节前第二天	0.1	-13.00	<0.01
节前第一天	0.1	-29.86	<0.01	节前第一天	0.1	-29.95	<0.01
节后第一天	0.1	13.52	<0.01	节后第一天	0.1	13.33	<0.01
节后第三天	0.1	5.91	0.097				

由表 2 可知经过多元回归分析,分别将节前第三天 $P=0.427>0.01$ 、节后第二天 $P=0.214>0.01$ 、节后第三天 $P=0.097>0.01$ 这 3 个显著性高的指标剔除后,得到星期、是否

一附属医院设立的我国唯一的国家级远程医疗中心,负责全国远程医疗技术发展研究、行业应用检测、技术标准和临床规范制定、行业交流等工作,也是中国卫生信息学会远程医疗专委会依托机构。

以国家远程医疗中心 2020 年 4—6 月远程会诊需求量为例,远程会诊需求量随日期变化如图 5 所示,首先,可以看出远程会诊需求量的变化具有明显的以周为周期的周期性,因此将星期 (周) 作为变量;其次,假期对需求量有明显影响。假期当天的需求量明显减少,并且假期前后的需求量也比普通工作日少,故假期长短和假期前后也对需求量有影响,为了将假期效应对远程会诊需求量的影响输入构建的预测模型中,对与假期相关的指标进行选择,将假期长短即假期前三天、前两天、前一天、后一天、后两天、后三天作为指标。

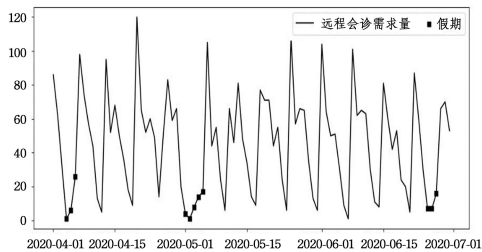


图 5 2020 年 4—6 月会诊量随日期变化

Fig. 5 The number of consultations from April to June 2020 changes with the date

使用多元回归分析对所选指标进行自变量和因变量间的相关性分析,每个自变量相关系数和显著性如表 2 所列。通过显著性大小可以判断每个指标对远程会诊需求量的影响程度,显著性越小表明该指标对远程会诊需求量影响程度越大,本研究取 $P<0.01$ 显著性,将显著性大于 0.01 的指标中显著性最大的指标剔除掉,在指标筛选过程中,一次只能剔除一个指标,之后对剩下的指标继续进行多元线性回归分析,经过多次分析得到的结果如表 2 所列。

节假日、节前第二天、节前第一天、节后第一天 5 个指标,将得到的 5 个假期相关指标和时间信息年、月、日共 8 个指标作为输入向量,如表 3 所列。

表 3 输入指标
Table 3 Input indicators

特征向量	指标	特征向量	指标
x_1	年	x_5	节前第二天
x_2	月	x_6	节前第一天
x_3	日	x_7	节假日
x_4	星期	x_8	节后第一天

3.1.3 数据归一化

在训练神经网络时,在不清楚各个指标的相对重要程度之前,数值大的指标对模型结果影响程度更大,为了消除指标之间的差异性,需要对数据进行标准化处理:

$$y_i = \frac{x_i - \bar{x}}{s}, i = 1, 2, \dots, t \quad (14)$$

其中, $\bar{x} = \frac{\sum_{i=1}^N x_i}{N}$, $s = \sqrt{\frac{(\sum_{i=1}^N x_i - \bar{x})^2}{N-1}}$, N 为指标数量。

数据归一化后如图 6 所示。

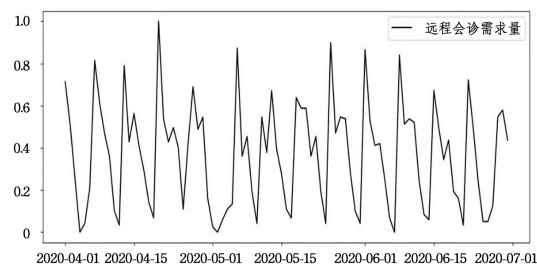


图 6 2020 年 4-6 月需求量归一化

Fig. 6 Demand normalization from April to June 2020

3.2 超参数设置

本节对模型调参:

1) LSTM 层数。一般设置为 1-2 层,经实验后得出模型层数设置为 2 时预测效果较好。

2) 学习率。本模型采用 Adam 算法,Adam 优化算法由 Kingma 和 Ba 提出^[19],Adam 在深度学习领域是一种很受欢迎的算法,它能很快取得好的成果。实证结果表明 Adam 的预测效果较好,并优于其他随机优化方法。Adam 参数设置:学习率为 0.001, beta1 为 0.9, beta2 为 0.999, epsilon 为 10^{-8} 。

3) 时间步长。由图 4 知远程会诊需求量存在以周为单位的周期性,故将时间步长设置为 7。

4) 批量大小。批量设置较大能提高内存利用率,提高计算效率和速度,但也易陷入局部最优,泛化性较差;批量设置较小时占用内存较小,训练速度更快同时有更好的泛化能力,但缺点是模型训练时间较长,同时因为每个 batch 的样本数目较少,可能不够代表整个数据集的分布,从而导致训练过程中出现较大的波动,使得模型训练不稳定。遍历[10, 20, 30, 40, 50]不同批量大小,模型预测结果如表 4 所列。

表 4 不同批量大小模型预测结果

Table 4 Model prediction results with different batch sizes

批量大小	RMSE	MAE	R2
10	11.769	8.709	0.808
20	12.066	8.332	0.799
30	13.444	9.585	0.750
40	14.633	9.868	0.704
50	14.741	10.160	0.699

由表 4 可知当批量大小设置为 10 时,模型预测效果最好。

5) 训练轮数。为了寻找合适的训练轮数,分别设置为 [50, 100, 150, 200, 250, 300] 进行多次训练,不同训练轮数对应训练集误差下降图如图 7 所示。

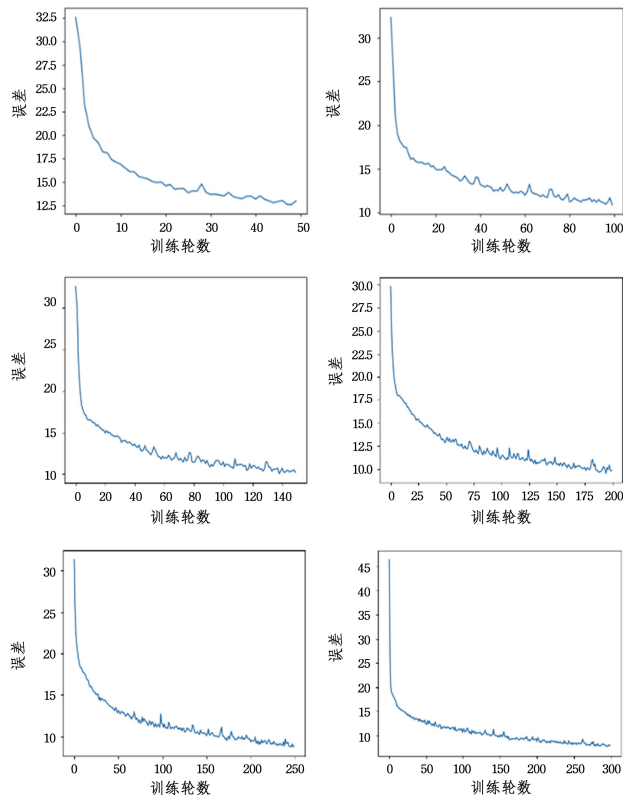


图 7 不同训练轮数训练集误差下降图

Fig. 7 Error drop diagram of training set with different number of training rounds

由图 7 可知,随着模型训练轮数不断增加,训练集误差不断下降,模型效果越来越好,误差在 0~50 个 epochs 快速下降,随后的 epochs 中误差下降速度逐渐变慢并逐渐收敛,约在 275 epochs 左右趋于稳定,故本实验模型训练轮数设置为 275。

6) Dropout。Dropout 指神经网络在向前传播的过程中,让某个神经元以一定概率停止工作,这样能使模型不太依赖某些局部特征,提高模型泛化性。经多次实验,Dropout 取 0.2 时模型拟合效果最好。

7) 神经元节点数。本文中两个 LSTM 层神经元数量设置相同,使用网格搜索算法确定 LSTM 层和注意力层神经元数量。神经元数量一般取 2 的次幂,故分别设置神经元数量为 [4, 8, 16, 32, 64] 进行交叉搜索寻优,不同神经元下模型预测效果如表 5 所列。

表 5 不同神经元下模型 RMSE

Table 5 Model RMSE under different neurons

注意力层 神经元数量	LSTM 层神经元数量				
	4	8	16	32	64
4	15.145	15.253	14.705	13.356	17.139
8	15.299	13.690	15.782	15.190	13.194
16	14.074	14.018	11.769	14.238	14.441
32	16.368	13.314	14.297	15.056	13.926
64	16.931	17.798	14.597	14.691	15.225

3.3 结果分析

3.3.1 假期指标适用性分析

由表 5 可知,当 LSTM 层神经元和注意力层神经元为 16 时,模型误差最低,预测效果最好。

为了探究假期指标在实证分析中的可行性和适用性,分别将带假期指标的数据和仅包含年、月、日、星期 4 项指标的原始数据输入到 MLR-Attention-LSTM 神经网络模型,其在测试集上的预测效果分别如图 8、图 9 所示,两种模型在测试集上误差如表 6 所列。

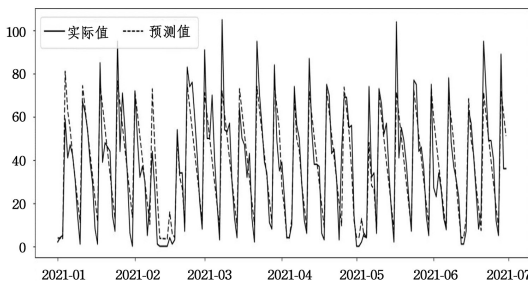


图 8 MLR-Attention-LSTM 神经网络模型预测值与实际值对比 (假期指标)

Fig. 8 Comparison of predicted value and actual value of MLR-Attention-LSTM neural network model (holiday indicator)

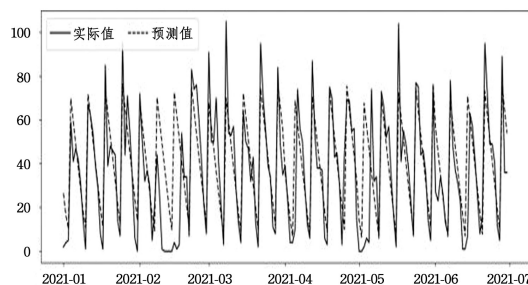


图 9 MLR-Attention-LSTM 神经网络模型预测值与实际值对比 (原始指标)

Fig. 9 Comparison of predicted value and actual value of MLR-Attention-LSTM neural network model (original index)

表 6 预测结果比较

	RMSE	MAE	R ²
原始指标	15.460	11.230	0.700
假期指标	11.769	8.709	0.808

由表 6 可得,输入假期指标的模型预测误差 RMSE 和 MAE 均低于使用原始指标的模型,由此可见假期信息的输入能有效降低预测误差,提高模型性能。

3.3.2 MLR-Attention-LSTM 与其他模型对比

为了验证 MLR-Attention-LSTM 神经网络的预测效果,选取 ARIMA, KNN, SVR, BPNN 和 LSTM 5 种模型作对比,并用 RMSE, MAE 和 R-squared 作为评估指标,不同模型预测结果如表 7 所列。

由表 7 可以看出,附加假期指标作为输入变量的模型测试集误差低于使用原始指标的模型,一方面说明考虑假期效应的 MLR-Attention-LSTM 神经网络预测效果更好,另一方面说明假期信息的输入可以提高模型的预测性能和泛化能力。

表 7 预测结果比较

Table 7 Comparison of forecast results

模型	测试集		
	RMSE	MAE	R ²
ARIMA	20.600	14.500	0.400
KNN	14.164	10.715	0.718
SVR	14.028	10.449	0.725
BPNN	12.937	9.745	0.760
LSTM	12.670	9.330	0.790
MLR-Attention-LSTM (原始指标)	15.460	11.230	0.700
MLR-Attention-LSTM (假期指标)	11.802	9.023	0.807

通过表 7 还可以看出,附带假期指标输入的 MLR-Attention-LSTM 神经网络在测试集上 RMSE 和 MAE 均为最低,表明其预测能力最好。BP 神经网络和 LSTM 神经网络由于其不能有效记忆数据中的有效信息,其模型复杂度低于 MLR-Attention-LSTM,故其预测精度略差于 LSTM,但仍比其他 3 种算法预测效果好。对于机器学习模型 KNN 和 SVR,可以看出 SVR 预测效果优于 KNN, KNN 算法的优点是简单,缺点是计算量大,尤其当数据维度较多时,当维度增加,数据点之间的距离越来越大,计算量大大增加的同时,对于 KNN 这种高度依赖距离的算法来说预测的准确率也会降低。SVR 通过核函数将样本从原始空间映射到一个更高维的特征空间,使得样本在这个特征空间内线性可分。通过对不同核函数测试可知在使用高斯核 rbf 时模型 RMSE 和 MAE 最小,模型预测效果最好。这两个机器模型预测效果和神经网络相比仍有差距,原因可能是其模型较为简单,容易出现过拟合,同时对输入向量存在的周期性和假期效应学习能力不足。传统的时间序列分析模型 ARIMA 的 RMSE 和 MAE 均为最高,原因可能是,ARIMA 是一种线性模型,无法学习远程会诊需求量与时间之间的非线性关系;其次由于 ARIMA 仅以历史数据作为输入,并没有捕捉到与远程会诊需求量相关的假期信息,而其他几个模型均能捕捉到输入数据所携带的假期信息,进而能更好地拟合数据来描述变量间的非线性关系。

总而言之,MLR-Attention-LSTM 神经网络具有记忆性,克服了 BP 神经不能有效保留数据中前刻的有效信息的缺点,从而能更好地捕捉数据之间的时间关系和序列信息。相较于传统时间序列模型,MLR-Attention-LSTM 克服了无法学习非线性关系的问题,同时也克服了机器学习容易出现过拟合和学习能力不足的缺点。同时,假期指标的输入进一步提高了模型的预测能力,证实了输入假期指标在远程会诊需求预测中的可行性与适用性。因此,相较于其他模型,考虑假期效应的 MLR-Attention-LSTM 神经网络能够更好地捕捉到输入数据中存在的非线性的周期性和假期效应,表现出了更好预测效果。

结束语 本文研究了 MLR-Attention-LSTM 神经网络对于远程会诊需求量预测的理论基础和实际应用,提出了远程会诊需求量存在的假期效应并选择假期相关指标作为模型输入,同时将 LSTM 神经网络模型和多元回归分析和注意力机制相结合,对模型的输入进行降噪并对隐藏层的输出进行分配权重,以此提高模型预测能力。基于国家远程医疗中心的实际会诊历史数据进行实证分析得到以下结论:

1)MLR-Attention-LSTM 神经网络相较于 BP 神经网络和 LSTM 神经网络预测更精确,可以在处理序列数据时保留前面时刻的信息,更好地捕捉数据之间的时间关系和序列信息,提高模型预测性能,因此模型本身复杂度影响模型的预测结果。深度学习等复杂度较高的模型因其本身拥有强大的学习能力,对数据的拟合效果更好,与注意力机制的结合使模型预测结果更加精准。

2)相较于 ARIMA,KNN,SVR 模型,LSTM 神经网络以其强大的对假期效应等非线性特征的学习能力,取得了更好的预测效果。预测结果可为国家远程医学中心优化资源配置提供理论支撑和决策参考。

深度学习模型在供应链预测、能源预测等许多领域取得巨大成功与广泛应用,但是在远程会诊领域的研究和应用相对较少。本文从理论上研究了深度学习的 LSTM 神经网络对远程会诊需求量预测的可行性,从大数据角度附加假期变量输入到模型中,并基于国家远程医疗中心的实际会诊数据证实了 LSTM 神经网络和考虑假期效应在远程会诊需求预测中的适用性。这对于构建大数据下远程会诊预测模型具有重要借鉴意义。

虽然通过构建 LSTM 神经网络模型和附加假期指标取得了不错的预测效果,但在相关变量的选择和模型的选择上仍值得进一步探索。本研究表明添加与需求相关的变量可以提高模型的预测精度,因此更多与需求相关的变量有待发现并导入到预测模型中。对于模型的选择,可以进一步探索更多深度模型或者组合模型。

参 考 文 献

[1] ZHAO J,CUI Z Y,CAI Y L,et al. Analysis on the Efficiency Optimization of Resource Allocation Based on Telemedicine[J]. Chinese Health Economics,2014,33(10):5-7.

[2] CHEN X,WANG J. Matching Method for Medical Service Supply and Demand Considering Bodies' Psychological Behavior Based on Intelligent Platform[J]. Operations Research and Management Science,2018,27(10):125-132.

[3] DONG T S,ZHANG M K. Practice and Thinking of Using Appointment Service of Hospital Outpatient in Telemedicine Consultation Scheduling Work[J]. Chinese Hospital Management, 2017,37(1):40-41.

[4] LU W,GAO P,ZHAI Y K. An Adaptive Recommendation Method for Telemedicine Specialists with Feedback Adjustment [J]. Journal of Systems & Management,2023,32(5):960-975.

[5] PANDA S K,MOHANTY S N. Time Series Forecasting and Modelling of Food Demand Supply Chain based on Regressors Analysis[J]. IEEE Access,2023,11:42679-42700.

[6] WANG Q,JIANG H,QIU M,et al. TGAE:Temporal Graph Autoencoder for Travel Forecasting[J]. IEEE Transactions on Intelligent Transportation Systems,2022,24(8):8529-8541.

[7] RUNGE J,ZMEUREANU R. Deep learning forecasting for

electric demand applications of cooling systems in buildings[J]. Advanced Engineering Informatics,2022,53:101674.

[8] XUE G,LIU S,REN L,et al. Forecasting hourly attraction tourist volume with search engine and social media data for decision support[J]. Information Processing & Management,2023, 60(4):103399.

[9] BUCKINGHAM-JEFFERY E,MORBNEY R,HOUSE T,et al. Correcting for day of the week and public holiday effects:improving a national daily syndromic surveillance service for detecting public health threats[J]. BMC Public Health,2017,17(1):1-9.

[10] OUYANG H B,HUANG K,YAN H J. Prediction of Financial Time Series Based on LSTM Neural Network[J]. Chinese Journal of Management Science,2020,28(4):27-35.

[11] CHEN W J,YU L,LI J L. Forecasting Teleconsultation Demand with an Ensemble Attention-Based Bidirectional Long Short-Term Memory Model[J]. International Journal of Computational Intelligence Systems,2021,14(1):821-833.

[12] TANG Z P,WU J C,ZAHNG T T,et al. An EEMD-LSTM Model Based Research on Early Warning of the Systematic Risk in China Insurance Industry[J]. Management Review, 2022, 34(9):27-34.

[13] LIU Y M,LI Y,ZHAO Z Y. Forecasting Price Trend of Constituent Stocks Using RF-LSTM Model Based on Feature Selection [J]. Statistics & Decision,2021,37(1):157-160.

[14] LIU S,YAO E. Holiday passenger flow forecasting based on the modified least-square support vector machine for the metro system[J]. Journal of Transportation Engineering, Part A-Systems,2017,143(2):1-8.

[15] YAO E,HONG J,PAN L,et al. Forecasting Passenger Flow Distribution on Holidays for Urban Rail Transit Based on Destination Choice Behavior Analysis [J]. Journal of Advanced Transportation,2021,2021:1-13.

[16] HOCHREITER S,SCHMIDHUBER J. Long Short-Term Memory[J]. Neural Computation,1997,9(8):1735-1780.

[17] KINGMA D P,BA J. Adam:A method for stochastic optimization[J]. arXiv:1412. 6980,2015.



ZHAI Yunkai, born in 1980, Ph.D, professor, Ph.D supervisor. His main research interests include healthcare big data and telemedicine information system and management.



QIAO Yan, born in 1991, Ph. D. His main research interests include healthcare big data and medical informatization.