

基于深度学习的清远市景区流量预测模型研究

李国平

广东碧桂园职业学院

DOI:10.12238/acair.v2i2.7373

[摘要] 为提高旅游景区流量预测的准确性,本文提出了一种基于深度学习的清远市旅游景区流量预测模型。首先,通过景区售票系统收集游客的进出景区信息。其次,对数据进行初步处理,包括数据清洗、归一化整理。然后,利用深度学习中的卷积神经网络算法构建了一种新型的旅游景区流量预测模型,该模型利用历史景区流量数据进行训练,并能够自动学习数据中的复杂特征和规律,从而精准预测未来景区流量变化。实验结果显示,所设计的模型精度达到87.61%,较传统方法提高了16.18个百分点,表明新模型在景区流量预测准确性方面取得了显著的提升。本文提出的基于深度学习的旅游景区流量预测模型,可为景区运营管理提供更可靠的决策支持,从而提高景区管理效率。

[关键词] 深度学习; 旅游景区; 流量预测; 卷积神经网络算法

中图分类号: F593/597 **文献标识码:** A

Research on Deep Learning Based Scenic Area Traffic Prediction Model in Qingyuan City

Guoping Li

Guangdong Country Garden Polytechnic

[Abstract] In order to improve the accuracy of traffic prediction for tourist attractions, this paper proposes a traffic prediction model for tourist attractions in Qingyuan City based on deep learning. Firstly, collect information on tourists entering and exiting the scenic area through the ticketing system. Secondly, perform preliminary data processing, including data cleaning and normalization. Then, a new type of tourist attraction traffic prediction model was constructed using convolutional neural network algorithms in deep learning. The model is trained using historical attraction traffic data and can automatically learn complex features and patterns in the data, thereby accurately predicting future changes in attraction traffic. The experimental results show that the designed model has an accuracy of 87.61%, which is 16.18 percentage points higher than traditional methods, indicating that the new model has achieved significant improvement in the accuracy of scenic area traffic prediction. The deep learning based tourism scenic area traffic prediction model proposed in this article can provide more reliable decision support for scenic area operation and management, thereby improving the efficiency of scenic area management.

[Key words] Deep learning; Tourist attractions; Traffic prediction; Convolutional neural network algorithm

引言

清远市作为著名的旅游城市,拥有众多具有代表性的旅游景区,吸引了大量游客前来游览。景区流量的精准预测对景区运营管理具有重要意义,直接影响到景区的安全管理、资源调度和服务质量等方面。因此,准确预测景区流量是当前景区管理中必不可少的一项工作。

然而,现有的预测方法在应对复杂的客流数据时存在一定局限性,特别是在处理非线性、非平稳和大规模数据时表现较差。鉴于此,本文提出了一种基于深度学习的景区流量预测模型,旨在利用深度学习技术在特征提取和模式识别方面的优势,实

现对景区流量的精确预测。相较于传统的预测方法,深度学习模型具有更强的自适应性和学习能力,可以更好地捕捉景区流量数据中的复杂特征和内在规律,为景区运营管理提供更可靠的决策支持,从而提高景区管理效率。

1 基于深度学习的景区流量预测模型

1.1 景区数据采集与预处理

景区客流数据的采集与预处理是构建流量预测模型的基础。首先,数据来源主要来源于景区售票系统。售票系统记录了游客的购票信息,包括门票类型、价格、景区有效时间、购票数量、游客信息等。在数据采集方面,需要设计合适的数据收集方式,并确保

数据的有效存储和管理。采集的原始数据可能包含重复、错误或异常的数据点,这些数据可能会对模型的训练和预测产生不利影响,需要进行预处理,包括数据清洗、归一化和转换等步骤^[2]。数据清洗是为了去除重复、错误和异常的数据;数据归一化是为了统一不同量纲的数据,使其在数值上具有可比性;数据转换则是将原始数据转换为适用于模型构建的格式。通过这些预处理步骤,为后续构建景区流量预测模型提供了高质量的数据基础。

将收集的数据需要进一步处理,经过归一化处理后得到预测的分析数据,数据转换计算公式为:

$$Y = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

公式(1)中Y表示归一化处理后的人流量的数, X_{\min} 表示数据集中最小的人流量数据, X_{\max} 是数据集中最大的人流量数据。 X 是原始收集的人流量数据值。通过公式(1)的处理、转换后,可以对采集的旅客信息进行初步处理,为后续的预测模型提供有效数据支撑。

1.2 基于深度学习的建立景区流量预测模型

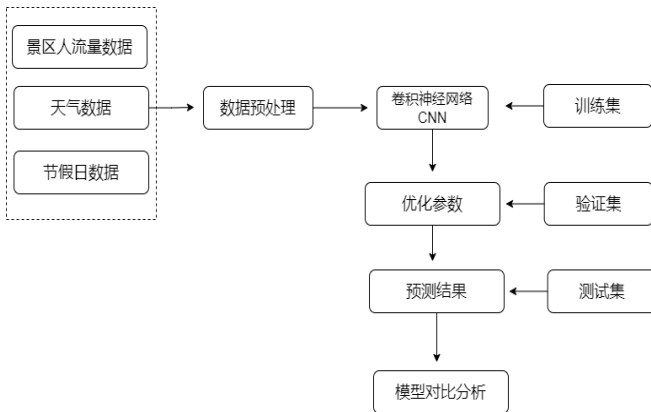


图1 景区人流量构建模型流程图

采用了卷积神经网络(CNN)作为基础模型架构,以利用其在时空数据特征提取方面的优势^[5]。基于CNN构建景区流量预测模型的具体流程如下:首先,设计CNN模型结构,确定模型架构,包括卷积层、池化层、全连接层等,以适应景区客流量的数据特点。其次,对数据进行预处理,对历史客流量数据进行处理,清除其中的重复条目、明显错误和不符合常规的数据点。接下来,实施数据标准化,以消除不同度量单位对分析结果的影响。在此基础上,再次进行模型的构建与训练环节,通过历史客流量数据对卷积神经网络(CNN)模型进行训练,旨在精确调整模型的参数配置。同时,运用验证集对模型的性能进行检验,并根据结果调整模型的结构和超参数设置,以期达到最优的预测效果。再次,在模型性能评价与提升阶段,采用诸如均方误差(MSE)和决定系数(R^2)等评估指标来衡量模型的准确性和效果。通过开展对比实验,检验并证实所构建模型的效能。最后,将经过充分训练与调优的模型应用于实际环境,实现对数据的即时预测与分析,根据预测结果,进行景区资源调度和安全管理。整体的流程图如图1所示。

在构建基于CNN的景区流量预测模型的过程中,为了缩小预测结果与实际观测值之间的偏差,选用均方误差(MSE)作为评价损失的指标,该指标有效地反映了预测误差的平方平均值,利用反向传播法对MSE进行计算,以优化并更新模型的参数。MSE的计算遵循以下公式:

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

公式(2)中MSE表示均方误差,n表示样本数量, y_i 表示真实值, \hat{y}_i 表示预测值。

在训练过程中,采用反向传播方法来确定每个参数的梯度,进而利用Adam优化器进行参数的调整。Adam优化器以其自适应的学习率调整机制而著称,能够在训练过程中根据参数的梯度及其变化趋势,自动地对每个参数的学习率进行动态优化。这种方法有助于加速模型向最小损失方向的收敛,提高模型训练的效率。该算法优化结果的计算公式为:

$$\alpha = \beta_1 g + \frac{\beta_2 g^2}{2} \quad (3)$$

公式(3)中 α 表示Adam优化结果, g 表示梯度, β_1 , β_2 分别表示学习率与衰减率。

采用交叉验证方法来评估基于CNN的景区流量预测模型的稳定性和泛化能力。首先,初始阶段,原始数据被分割成三个部分:训练集、验证集和测试集。训练集用于构建和训练模型,验证集用于微调模型架构和超参数设置,而测试集则用来检验模型在未见过的数据上的泛化能力。其次,在模型训练过程中,使用训练集数据来调整网络结构和超参数。同时,利用验证集对模型进行性能评估,这有助于预防模型过度拟合训练数据。然后,我们使用测试集数据来评估模型的泛化能力。具体而言,通过计算诸如准确率、召回率和F1分数等评估指标,我们可以对模型的预测准确性与实际值之间的偏差进行量化分析,从而评估模型的预测性能。最后,依据测试集的反馈,对模型架构和超参数进行了优化调整,并循环执行模型的训练、评估与改进流程。经过多次迭代和优化后,使用测试集数据评估了模型的性能^[4]。

对基于CNN的景区流量预测模型的网络结构和超参数进行了细致的调整^[3]。首先,根据模型的预测性能和交叉验证的结果,对当前网络结构的性能进行了评估。接着,基于评估反馈,对网络架构进行了优化,包括增加或减少层数以及改变层的大小等。最后,对超参数进行了优化,包括调整学习率、批量大小和迭代轮数等,以进一步提高模型的整体性能。通过这些调整,优化了模型,提高了其预测精度。

2 实验论证

为了观察模型预测旅客人流量与真实旅客之间的误差,将预测的结果绘制成曲线图,通过潘念然(2022)文献中的ARIMA模型^[1](自回归积分滑动平均模型)作为对照组,如图2所示,与本研究设计的模型预测曲线图进行对比,如图3所示。

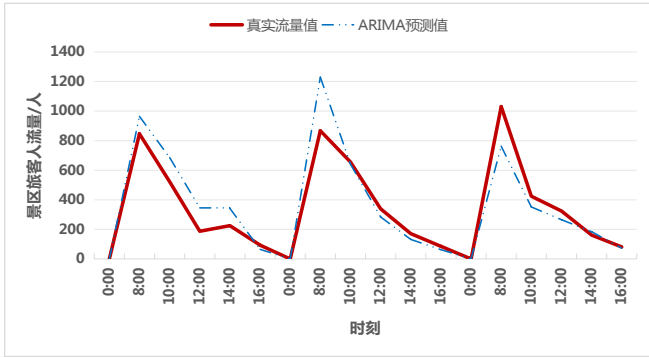


图2 ARIMA模型预测值与真实流量值对比图

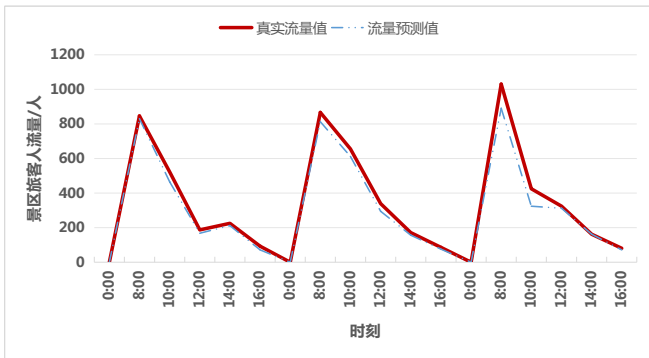


图3 本设计模型与真实流量值对比图

通过图2与图3可知,直观可以看出整体曲线的趋势大同小异,但图2可看出,在较多的时间段内游客人流量与预测值相差较大,可通过均方根误差进行计算,其计算公式为:

$$R_{MSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (4)$$

公式(4)中 R_{MSE} 表示均方根误差, n 表示样本数据, y_i 表示真实流量值, \hat{y}_i 表示预测值。

表1 不同预测模型均方根误差对比

模型	均方根误差值	预测精准度(%)
本研究构建模型	21.3346	87.61
ARIMA 模型	32.5781	71.43

通过公式(2)分别计算ARIMA模型预测与本研究模型预测的均方根误差,得出结果如表1所示。

3 结语

本研究以清远市旅游景区为研究对象,提出了一种基于深度学习的景区流量预测模型。通过对景区售票系统收集的客流数据进行预处理,我们利用卷积神经网络(CNN)技术构建了预测模型。通过实验验证,我们发现该模型能够有效地从历史数据中学习并预测未来景区流量变化,预测精度达到87.61%,较传统方法提高了16.18个百分点。

本研究的创新之处在于:

(1)利用深度学习技术,特别是卷积神经网络(CNN),对景区流量数据进行特征提取和模式识别,提高了预测的准确性。(2)通过对数据进行预处理,包括清洗、归一化等步骤,提高了模型的泛化能力。(3)模型在预测精度上显著优于传统方法,为景区运营管理提供了有力的决策支持。

本研究的结果表明,基于深度学习的景区流量预测模型具有较好的稳定性和泛化能力,可以为景区管理提供可靠的数据支持。未来,我们将继续优化模型,提高预测精度,并探索更多的数据源,以进一步提升景区流量的预测准确性。此外,我们还将研究基于深度学习的景区资源调度和安全管理策略,以实现景区的可持续发展。

[基金]

基于深度学习的清远市旅游景区流量预测模型研究(QYSK2023048)。

[参考文献]

- [1]潘念然.基于ARIMA和LSTM的城市轨道交通客流量预测[J].科学技术创新,2022,(08):165-168.
- [2]刘帆,朱强.基于深度学习的轨道交通客流量预测模型[J].信息与电脑(理论版),2023,35(22):63-65.
- [3]李亚香,王元庆.基于PSO-LSSVM算法的城市轨道交通断面客流量预测[J].长安大学学报(自然科学版),2021,41(6):91-102.
- [4]谢谦,陆明,谢春山.基于LBS和深度学习的旅游景区客流量的高时频预测[J].地球信息科学学报,2023,25(02):298-310.
- [5]李安娜,李永福.基于深度神经网络的短时地铁客流预测[J].运筹与模糊学,2022,12(1):26-35.