

基于CNN-GRU-ATT的气温预测算法研究

杜帅文^{1,2}, 张凯钧³, 刘霄燕^{1,2}, 王双双^{1,2}, 翁秀年^{1,2}

(1. 海南省气象信息中心; 2. 海南省南海气象防灾减灾重点实验室, 海南海口 570203;
3. 海南警察学院, 海南海口 570100)

摘要: 当前气温预测在气象预报中具有重要意义, 特别是对于应对极端天气和优化各行业的决策至关重要, 准确的气温预测对于城市的经济发展和人类社会活动不可或缺。以海口市为例, 提出一种基于深度学习的气温短期预测算法, 采用CNN-GRU-ATT模型架构以提高传统模型气温预测精度。首先, 利用Pearson相关系数法对海口气象相关数据进行特征选择, 筛选出的特征作为模型输入; 其次, 通过卷积神经网络(CNN)提取相关局部特征, 捕捉多维历史数据中的有效信息; 再次, 模型进一步采用门控循环单元(GRU)对特征序列进行时间序列建模, 以识别气温变化的依赖关系; 最后, 引入注意力机制(Attention)对GRU的输出进行加权, 使模型能够集中关注重要的时间步, 忽略不相关信息, 从而优化预测效果。在2005—2024年海口市真实气温相关数据集上的测试结果显示, 该算法在RMSE和MAE两个评价指标上具有显著优势, 其中RMSE为2.21, MAE为1.65, 优于其他对比模型预测结果, 为海口市气温短期预测提供了有效的参考方案。

关键词: 深度学习; 气温预测; 卷积神经网络; 门控循环单元; 注意力机制

DOI: 10.11907/rj.241911

扫描二维码阅读全文:



中图分类号: TP391

文献标识码: A

文章编号: 1672-7800(2026)002-0083-08

Research on Temperature Prediction Algorithm Based on CNN-GRU-ATT

DU Shuaiwen^{1,2}, ZHANG Kaijun³, LIU Xiaoyan^{1,2}, WANG Shuangshuang^{1,2}, WENG Xiunian^{1,2}

(1. Meteorological Information Center of Hainan Province;

2. Key Laboratory of South China Sea Meteorological Disaster Prevention and Mitigation of Hainan Province, Haikou 570203, China;

3. Hainan Police College, Haikou 570100, China)

Abstract: Temperature prediction is of great significance in weather forecasting, especially for dealing with extreme weather and optimizing decision-making in various industries. Accurate temperature prediction is indispensable for the economic development of cities and human social activities. Taking Haikou City as an example, this paper proposes a short-term temperature prediction algorithm based on deep learning, and adopts the CNN-GRU-ATT model architecture to improve the accuracy of the traditional temperature prediction model. First, the Pearson correlation coefficient method is used to select features of Haikou meteorological data, and the selected features are used as the input of the model. Secondly, the convolutional neural network (CNN) is used to extract relevant local features to capture effective information in multidimensional historical data. The model further uses the gated recurrent unit (GRU) to model the feature sequence in time series to identify the dependency of temperature changes. Finally, the attention mechanism is introduced to weight the output of GRU, so that the model can focus on important time steps and ignore irrelevant information, thereby optimizing the prediction effect. The test results on the real temperature related dataset in Haikou from 2005 to 2024 show that the proposed algorithm has significant advantages in the prediction accuracy of RMSE and MAE, in which the RMSE is 2.21 and the MAE is 1.65, which is better than the prediction results of other comparison models, and provides an effective reference scheme for the short-term prediction of temperature in Haikou.

Key Words: deep learning; temperature prediction; convolutional neural network; gated recurrent unit; attention mechanism

收稿日期: 2024-12-03

基金项目: 海南省自然科学基金面上项目(623MS118)

作者简介: 杜帅文(2000-), 男, 硕士, 海南省气象信息中心、海南省南海气象防灾减灾重点实验室助理工程师, 研究方向为气象数据处理与应用; 张凯钧(1999-), 男, 硕士, 海南警察学院助教, 研究方向为机器学习、网络安全。

0 引言

自 21 世纪以来,天气变化对人类的日常生活和社会经济发展产生了显著影响,而通过提前数日的天气预报,人们能够更有效地安排未来的活动和日程。气温预测在气象预报中占有重要地位,它不仅对农业、交通、能源等领域有重要影响,还为人们的日常生活提供了便利。

近年来,气象数据的规模和维度剧增,更是对气象预测提出了新的挑战^[1]。因此,通过挖掘历史气象数据提前预测未来城市气温变化趋势,可以极大提高人们的生活品质和社会经济水平。

气温预测是一个时间序列问题,通常由多变量特征组成,因此可以利用相应的预测模型以估计未来气温变化^[2]。由于时间序列数据模式的季节性变化和不规则趋势,传统的机器学习技术无法学习数据序列模式以进行准确预测。传统的气温预测通常采用基于统计的技术分析方法,而忽略了其他相关气象特征与气温的关联性。时间序列分析方法如 ARIMA (Auto-Regressive Integrated Moving Average) 模型在气温预测中应用广泛,这类模型通过分析气温数据的时间序列特性,建立数学模型进行预测^[3]。然而,ARIMA 模型在处理非线性数据和多变量数据时存在局限性,其需要数据满足一定的约束条件才能使用,而且该方法也因准确性和稳定性不佳而受到诟病。近年来,人工智能,特别是深度学习技术的快速发展,为气温预测提供了新的解决方案,越来越多的研究者开始将其应用于气温预测^[4]。相比传统的统计学方法,深度学习模型具有强大的特征提取和模式识别能力,能够在大数据环境下自动学习和提取复杂的非线性关系,从而显著提高预测准确性^[5]。深度学习方法如长短期记忆网络(Long Short-Term Memory, LSTM)、卷积神经网络(Convolutional Neural Networks, CNN)、循环神经网络(Recurrent Neural Network, RNN)因其强大的特征提取能力和非线性拟合能力,逐渐成为气温预测研究的热点。为了克服循环神经网络(RNN)在长距离记忆中所带来的问题, Hochreiter 等^[6]、Gers 等^[7]提出将 LSTM 作为 RNN 的改进结构, LSTM 在 RNN 的基础上引入了门机制以更新信息,但 LSTM 结构复杂,训练和预测的时间成本也较高。Huang 等^[8]进一步优化了 LSTM 的结构,提出基于门控循环单元(GRU)的神经网络,将原来 LSTM 中的门控单元数量从 3 个减少到 2 个,简化了 LSTM 的计算量。目前,基于深度学习建模的气象预测模型大多都属于单一模型,这种单一模型并不能全面地抽取数据特征,因此使得最终预测误差偏大^[9]。合理的特征表达对于提高预测精度非常重要,基于 CNN-LSTM 的混合神经网络模型已广泛应用于各种时间序列预测,并已成为最流行的基于回归的预测方法。Sajjad 等^[10]提出一种混合 CNN-GRU 短期用电量预测模型,与 CNN、LSTM 和

CNN-LSTM 模型相比,该模型误差最小,这种组合受益于 CNN 强大的特征提取能力和 GRU 强大的时间特征学习能力,通常可以获得更高的预测精度。

一个城市的气温变化与多个特征相关联,例如相对湿度、风速、降雨量等气象特征。海口市作为中国南部海南省的省会城市,位于北纬 $19^{\circ}31'$ ~ $20^{\circ}04'$, 东经 $110^{\circ}07'$ ~ $110^{\circ}42'$ 之间,其地理位置特殊,处于热带地区,气候变化复杂多变,季风气候显著,受海洋影响较大,常年高温高湿,气温变化频繁且幅度较大,因而准确的气温预测对于海口市的防灾减灾、农业生产、旅游业等具有重要意义。然而,传统的气温预测方法在面对海口如此复杂的气候系统时,往往预测精度不足。

综上,本文提出一种基于深度学习的气温短期预测算法,通过融合多种气象变量,构建一个高效的深度神经网络模型。首先,收集海口市过去 20 年的历史气温数据及相关气象数据,进行详细的数据预处理和特征选择,对相关数据集进行缺失值填充以及归一化,并使用 Pearson 相关系数法对预处理过后的数据集进行特征选择。其次,设计一个 CNN-GRU-ATT 预测模型网络模型,用于捕捉气温变化的时间序列特征,该模型其主要由 CNN、GRU、Attention 机制 3 个框架组成,采用 CNN 对历史气温数据提取局部特征和空间特征,使用多层 GRU 处理序列数据并捕捉时间上的依赖关系,并结合 Attention 机制实现了对海口市气温的短期预测,这些模块的结合使得模型能够更好地从序列数据中提取特征、利用时间信息,并关注对预测结果有重要贡献的部分,提高了预测准确性和可靠性。最后,通过与其他预测模型进行对比实验及分析,实验结果证明了本文算法的优越性。本文研究为海口市气温预测提供了一种新的思路和方法,对于提高海口市气温短期预测准确性具有重要的理论和实践意义,同时也为其他类似地区的气温预测研究提供了有益参考。

1 相关理论

1.1 卷积神经网络模型

卷积神经网络(Convolutional Neural Networks, CNN)是当前比较常见的深度学习模型,近年来广泛应用于时间序列数据处理。其通过卷积操作能够有效提取局部特征,并具有参数共享和稀疏连接的优势,从而大幅减少模型参数数量^[11]。CNN 对网格数据具有强大的处理能力,其主要结构为卷积层、池化层和全连接层。卷积层是 CNN 的核心,通过卷积核在输入数据上进行滑动窗口操作,从而提取局部特征。对于时间序列数据,卷积操作可以捕捉到数据的局部模式和趋势。而池化层通常跟随在卷积层之后,用于下采样特征图,可以减少数据的维度和计算量,同时保留重要特征。全连接层将卷积层和池化层提取的特征进行整合,用于最终的分类或回归任务。CNN 基本结构如

图1所示。

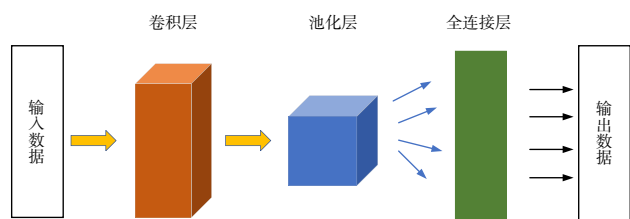


Fig. 1 CNN structure

图1 CNN结构

在气温预测任务中,CNN通过卷积层提取气象数据的局部时空特征,再通过池化层进行特征的降维和聚合,有助于捕捉到气温变化的局部模式和趋势,其公式如式(1)所示。

$$h_i = ELU(W_i \odot X + b_i) \quad (1)$$

其中, ELU 是激活函数, W_i 是权重矩阵, $\odot$ 表示卷积运算, b_i 表示偏置向量, h_i 是卷积层的输出结果, x 为输入数据。

1.2 门控循环单元模型

门控循环单元(Gated Recurrent Unit,GRU)是循环神经网络(RNN)的一种改进形式,专门用于处理序列数据。与标准RNN相比,GRU引入了门控机制,通过重置门(Reset Gate)和更新门(Update Gate)实现对信息流动的控制^[8]。这能够更好地捕捉长期依赖关系,并缓解梯度消失问题。更新门决定了前一时刻的状态信息在多大程度上保留到当前时刻,其输出介于0~1区间,通过与前一时刻的状态信息相乘,以控制信息流动。重置门决定了前一时刻的状态信息在生成候选状态时的影响程度,其输出也是0~1区间,通过与前一时刻的状态信息相乘,以调节对候选状态的影响^[12]。GRU结构如图2所示。

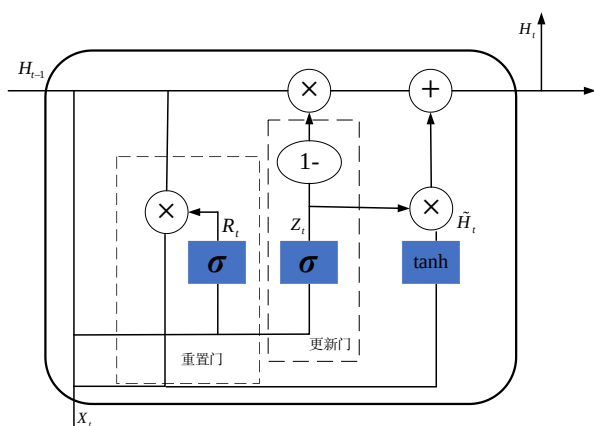


Fig. 2 GRU structure

图2 GRU结构

GRU通过更新门和重置门的协同作用,有效捕捉时间序列数据的长期依赖关系。在气温预测任务中,GRU可以对长时间的气温变化进行建模,并提供准确的短期预测结果。GRU计算过程如式(2)、式(3)所示。

$$Z_t = s(X_t W_{xz} + H_{t-1} W_{hz} + b_z) \quad (2)$$

$$R_t = s(X_t W_{rx} + H_{t-1} W_{rx} + b_r) \quad (3)$$

Z_t 代表更新门的输出, R_t 代表重置门的输出, σ 代表Sigmoid函数, W 代表矩阵的权重, b 代表偏置向量^[13]。 H_{t-1} 表示 $t-1$ 状态下的输出。GRU的连续性由候选隐藏状态和隐藏状态组成,它们的具体方程如式(4)、式(5)所示。

$$\tilde{H} = \tanh(x_t W_{hx} + R_t \odot H_{t-1} W_{hh} + b_h) \quad (4)$$

$$H_t = (1 - Z_t) \odot H_{t-1} + Z_t \odot \tilde{H}_t \quad (5)$$

$\tilde{H}$ 代表候选隐藏层状态, H_t 代表 $t-1$ 状态下的输出, $\tanh$ 表示双曲正切激活函数, $\odot$ 表示点乘运算。如果 R_t 的值接近1,意味着之前的隐藏状态更重要,需要保留隐藏状态中更多的信息。当 R_t 的值接近0,意味着前面的隐藏层不重要,会被遗忘。 Z_t 代表更新门。同理,当 Z_t 的值逼近1, H_t 基本上等于 H_{t-1} ,这意味着状态没有更新。当 Z_t 的值接近0时, H_t 的值由 $\tilde{H}_t$ 决定。

1.3 注意力机制

注意力机制(Attention Mechanism)最早应用于机器翻译任务,用于解决长序列信息处理问题,其通过为序列中的每个元素分配不同的权重,突出重要信息,忽略次要信息,从而提高模型性能^[14]。注意力在增加神经元之间的连接方面具有很好的特点,因为它可以自适应地选择相关的隐藏状态并在每个时间步维持所有隐藏状态。本文在气温预测任务中,通过将注意力机制集成到GRU模型中,对GRU生成的隐藏状态进行动态权重调整,模型得以集中处理关键时间步的信息,同时忽略不相关数据,从而增强了对特征序列中关键部分的关注度,并进一步提升了预测精度和模型整体性能。

在注意力机制中,输入序列的每一个元素首先通过一个线性变换生成3个向量:查询向量(Query)、键向量(Key)和值向量(Value)。注意力得分是通过查询向量和键向量的点积计算而得到,该得分代表了查询向量与对应键向量的相似程度^[15]。之后,使用Softmax函数对这些得分进行归一化处理,以获得注意力权重。最后,用这些权重对值向量进行加权求和,得到当前时间步的输出^[16]。整个过程可用式(6)表示。

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (6)$$

其中, Q 代表查询向量矩阵, K 代表键向量矩阵, V 代表值向量矩阵, d_k 是键向量的维度。 $\sqrt{d_k}$ 是一个缩放因子,用于防止点积结果过大,进而影响Softmax函数的梯度。

2 模型设计

CNN-GRU-ATT模型完整的预测过程如图3所示。模型总体流程为:①对原始数据进行预处理,具体操作包括处理缺失值、数据归一化,该步骤旨在确保数据质量,使其适合后续的特征选择和模型训练;②使用皮尔逊相关系数

法对数据中的特征进行筛选,通过计算各特征与目标特征的相关性,选择对预测效果有显著贡献的特征,并移除噪声或无关特征;③将处理后的数据集分为训练集和测试集,训练集用于训练模型,测试集用于验证模型性能;④利用筛选后的特征,构建基于卷积神经网络(CNN)、门控循环单元(GRU)和注意力机制(ATT)的深度学习模型。CNN用于提取局部特征,GRU用于捕捉时间序列信息,而注意力机制则用于提高模型对重要特征的关注度;⑤使用训练集数据对模型进行训练,通过多次迭代,优化模型参数,使其能够准确预测气温。在测试集上进行预测后,将预测结果进行反归一化处理,使其恢复到原始尺度,以便于结果解释和实际应用;⑥通过对比预测结果与实际结果,评估模型预测性能。

2.1 数据分析与预处理

2.1.1 数据分析

本文使用的数据来自于 <https://rp5.ru> 网站,该网站包含了 241 个国家 2005—2024 年期间的气象相关数据。从中选取海口市(气象站号码:59758)2005—2024 年期间的气温数据和多种气象变量相关数据,如相对湿度、平均风速、降水量、水平大气压、过去 12 小时内最低气温等,共计 15 条相关的气象记录,总计 45 389 条数据量,每天数据采集间隔时间为 3 小时。

根据相关数据,针对海口市近 20 年以来的气温变化趋势绘制了一个可视化曲线图,如图 4 所示。可以看出,海口的大气温度呈现出比较明显的周期性波动,每年都有一个比较一致的温度变化模式,这也反映了海口的一个气

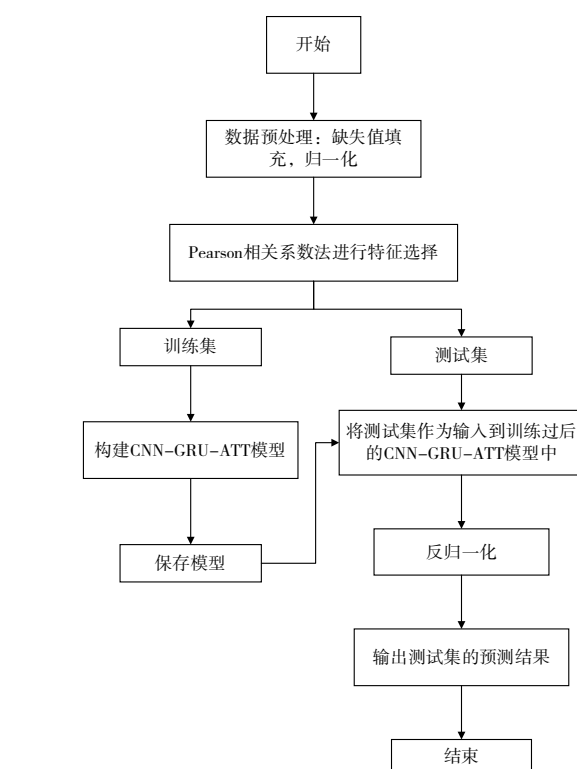


Fig. 3 Overall flow of temperature prediction

图3 气温预测总体流程

候特征,即通常在夏季达到最高温度,而在冬季温度最低,其中最低气温在5~10℃左右波动,最高气温在35~40℃左右波动,最高气温相对于最低气温持续时间较长,并且相邻时间段的气温差距较小。

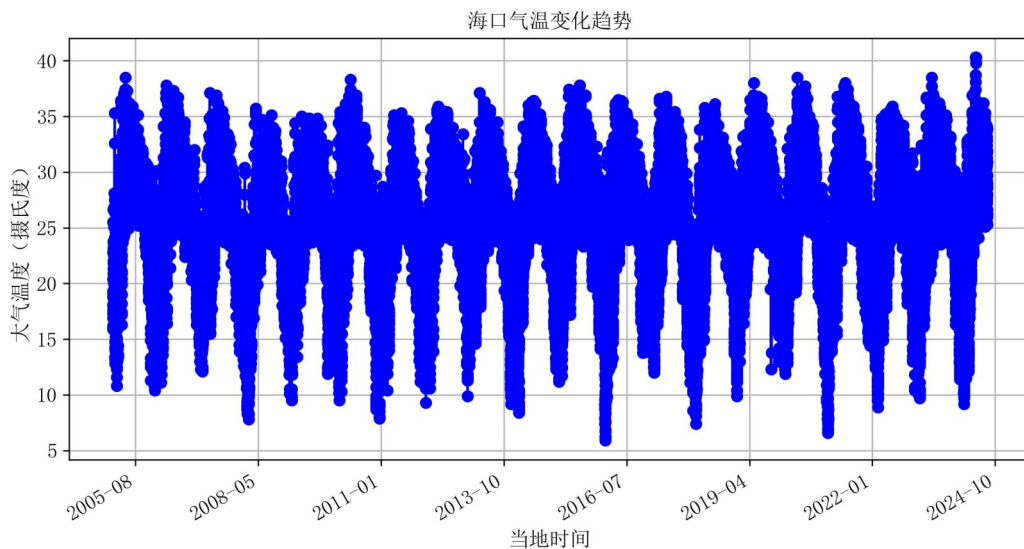


Fig. 4 Haikou temperature change trend

图4 海口气温变化趋势

2.1.2 数据预处理

在长期气象记录中,缺失值不可避免。对于少量随机缺失值,使用相邻时刻的均值进行填补。由于不同气象变量的量纲和取值范围不同,直接使用这些数据进行训练可

能导致模型性能下降,故本文对所有气象相关变量进行了归一化处理,将每个气象变量的值归一化到 $[-1, 1]$ 区间,使得所有气象变量的数据都在同一量纲范围内,以减少异常值的影响,从而加快模型收敛速度,提高训练效果。其

公式如(7)所示。

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (7)$$

其中, x 为原始数据, $x_{\min}$ 和 $x_{\max}$ 分别为该变量的最小值和最大值, x' 表示归一化后的数据。为了使预测结果具有现实意义,需要将预测结果进行反归一化处理,其公式如式(8)所示。

$$x = (x_{\max} - x_{\min})x' + x_{\min} \quad (8)$$

2.2 特征选择

特征选择在深度学习数据预处理中占据着关键位置,其主要任务是筛选出对模型预测性能有贡献的特征,并去除那些对预测结果影响不大或属于噪声的特征。这样不仅可以提升模型预测效率,还能够减少模型训练时的计算开销,并增强模型可解释性。本文将采用皮尔逊相关系数法以筛选影响气温的有效特征, Pearson 相关系数法属于一种过滤式特征选择方法,主要衡量特征与标签之间的相关性^[17]。

Pearson 相关系数法的步骤为:首先,针对数据集每个特征,计算其与目标变量之间的 Pearson 相关系数,以

评估它们之间的线性相关程度;其次,根据相关系数的值(范围为-1~1区间),评估特征与目标变量之间的相关性;再次,设定一个阈值以决定哪些特征的相关性足够强,保留那些相关系数的绝对值高于设定阈值的特征。通过这种方法,可以选出相关性较强且重要性分数高的特征作为最终输入变量。去除不重要的特征,不仅有助于降低模型复杂性,还能减少过拟合的可能性,从而提升模型泛化能力。其计算公式如式(9)所示。

$$r = \frac{n \sum xy - \sum x \sum y}{\sqrt{n \sum x^2 - (\sum x)^2} \sqrt{n \sum y^2 - (\sum y)^2}} \quad (9)$$

其中, r 为 Pearson 相关系数,取值范围为 $[-1, 1]$, 表示两个变量之间的相关程度。当 $r=1$ 时,表示两个变量完全正相关;当 $r=-1$ 时,表示两个变量负相关;当 $r=0$ 时,表示两个变量之间没有线性相关性。当 r 的绝对值越接近 1, 表示线性关系越强,绝对值越接近 0, 表示线性关系越弱。 n 表示样本数量; x 和 y 分别表示特征变量和标签变量。使用 Pearson 相关系数法对输入特征显示其特征重要性评分如图 5 所示。

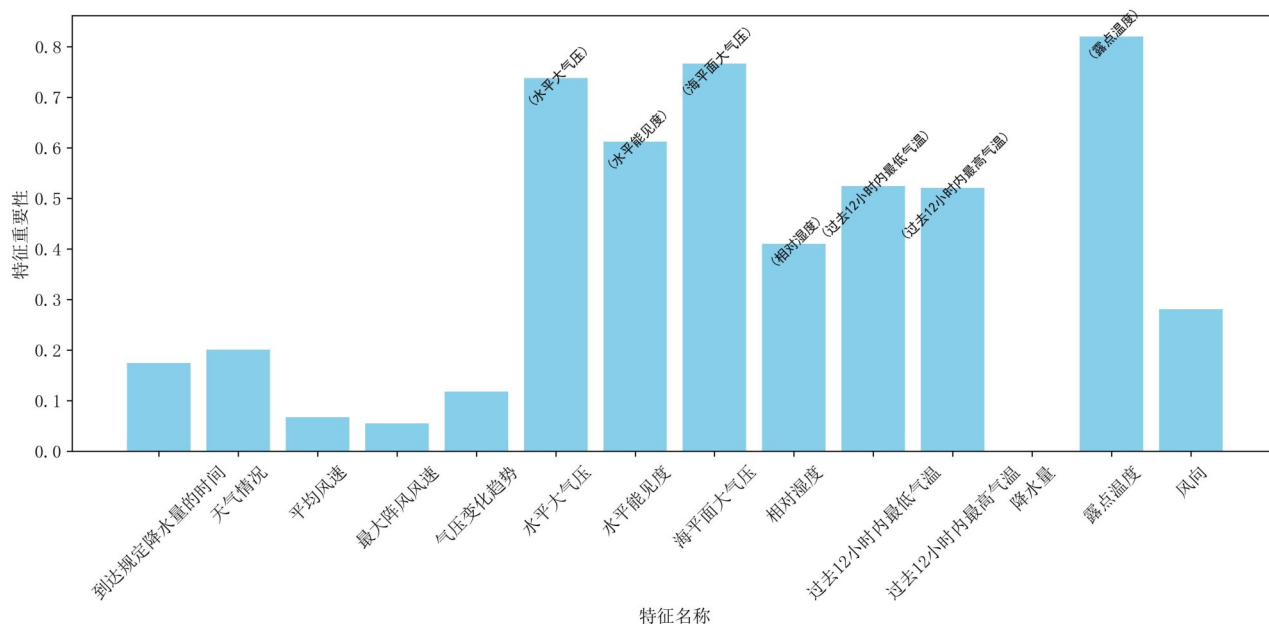


Fig. 5 Feature importance score

图5 特征重要性评分

为了充分考虑特征多样性和尽可能降低模型复杂度,本文将 Pearson 相关系数法的阈值设置为 0.4,即选择阈值大于 0.4 的特征,故筛选出温度、水平大气压、水平能见度、海平面大气压、相对湿度、过去 12 小时内最低气温、过去 12 小时内最高气温、露点温度这 8 个特征作为模型输入^[18]。为了训练和评估模型,本文将特征选择后的海口市气温相关数据集按 8:2 划分为训练集、测试集。

2.3 模型架构

本文提出的 CNN-GRU-ATT 融合模型由输入层、CNN

层、GRU 层、注意力层和输出层构成,整体结构如图 6 所示^[19]。模型以海口市的气温数据为输入,首先,通过 CNN 层进行特征提取,利用卷积操作捕捉数据项的空间特征及其内在联系;其次,通过池化操作以缩减模型参数规模,实现数据降维;再次,采用多层 GRU 进行有效的序列学习,GRU 层接收来自 CNN 层输出的特征图,并生成包含时间依赖性的信息流,这使得模型能够预测未来气温变化。在注意力层中根据 GRU 层的输出,计算每个时间步的重要性权重,进而聚焦于对预测最有贡献的特征;最后,

通过全连接层将注意力层的输出加以整合,用以最终气温

预测,最终输出未来几天的气温预测值。

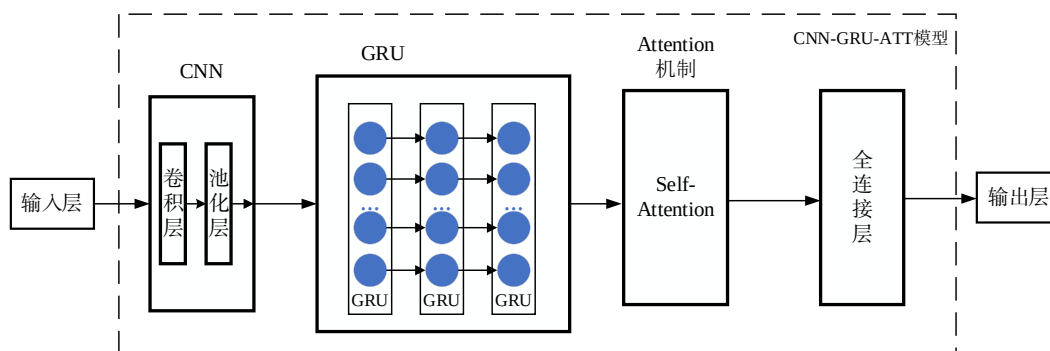


Fig. 6 CNN-GRU-ATT model architecture

图6 CNN-GRU-ATT模型架构

3 实验与结果分析

3.1 实验环境

本文实验硬件环境为 AMD R7-5800H 处理器, 16 GB 内存, 显卡为 NVIDIA GeForce RTX 3060 GPU。软件环境为 Windows 11 操作系统, PyCharm 编译器, 编程语言为 Python 3.9, 深度学习计算框架为 Pytorch 1.12.0。

3.2 评价指标

为了评估不同方法的预测性能, 评估指标选用均方根误差 (Root Mean Squared Error, RMSE) 和平均绝对误差 (Mean Absolute Error, MAE)。其中, MAE 表示模型预测值与真实值之间绝对误差的平均值, 用以反映预测值与真实值之间的误差^[4]; RMSE 表示模型预测值与真实值的平方误差均值的平方根, 用以衡量预测值与真实值之间的偏差^[20]。真实值与预测值的完美拟合会使 RMSE 和 MAE 为 0, 其值越接近 0, 代表模型预测效果越好; 其值越大, 代表模型预测效果越差。RMSE 和 MAE 计算公式分别如式 (10)、式 (11) 所示。

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (y(i) - \hat{y}(i))^2} \quad (10)$$

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

其中, T 代表测试样本总数, $y(i)$ 和 $\hat{y}(i)$ 分别代表真实值和预测值。

3.3 实验参数设置

在构建 CNN-GRU-ATT 预测模型时, 参数配置如下: 卷积层数量设为 1, 卷积核数量设定为 52 个, 尺寸为 5, 输入通道数设为单通道, 卷积核步长设为 1; 采用 3 个 GRU 层, 每层 GRU 的神经元数量为 39。针对具有显著时间序列特性的气象数据, 模型训练阶段采用 MSE 作为损失函数以优化性能。选定的特征维度是基于特征选择后的 8 维, 学习率设定为 0.000 1, 同时为了预防过拟合, Dropout 比率设置为 0.2。模型训练时, 迭代轮数定为 200 次, 每批次处理的样本数 (batch_size) 为 32。为了加快学习速度和提高

模型收敛效率, 实验采用了 Adam 优化算法, 选用 Relu 激活函数。

3.4 实验对比分析

3.4.1 模型性能对比

为了验证 CNN-GRU-ATT 组合模型对气温相关数据的预测性能以及模块的有效性, 本文实验中, 在相同参数条件下进行单步气温预测^[21]。单步预测是指基于时间序列中已经观察到的数据点, 对序列接下来一个时间点的值进行预测。由于本文所使用的数据集是每 3 小时记录一次数据, 因而实验中的单步预测即预测接下来 3 小时的气温数据。为了验证所提出模型的有效性和适用性, 选取 4 个模型 (GRU、CNN、LSTM、CNN-GRU) 作为对比模型进行分析, 结果如表 1 所示。

Table 1 Comparison of model prediction results

表 1 模型预测结果比较

模型名称	RMSE	MAE
CNN	2.85	2.24
LSTM	2.64	2.07
GRU	2.59	1.98
CNN-GRU	2.36	1.73
CNN-GRU-ATT	2.21	1.65

可以看出, 相较于 CNN、LSTM、GRU、CNN-GRU 模型, 本文 CNN-GRU-ATT 模型在 RMSE 和 MAE 评价指标上均低于以上对比模型, 表明其预测效果更好, 证明了该模型的优越性。其中, 单一模型 CNN、LSTM、GRU 相较于混合模型 CNN-GRU、CNN-GRU-ATT, 单一模型预测效果较差, 表明这些模型缺乏有效的组合优化, 导致其预测性能没有得到充分展示。CNN-GRU-ATT 模型在 CNN-GRU 模型的基础上增加了注意力机制, 使得让模型在每个时间步上关注更重要的特征, 从而非常有助于复杂时序数据处理, 故 CNN-GRU-ATT 相较于 CNN-GRU 模型预测效果更优越。CNN-GRU 模型在 CNN 模块的基础上增加了 GRU 模块, 使得 CNN-GRU 模型更关注时间序列信息, 故 CNN-GRU 模型相较于 CNN-GRU 模型预测效果更优越。

3.4.2 模型预测结果

为了展示本文 CNN-GRU-ATT 模型在训练集和测试

集上的直观预测效果,实验中绘制了时间步长为 1 的预测气温值与真实气温值对比曲线,如图 7 所示^[22]。

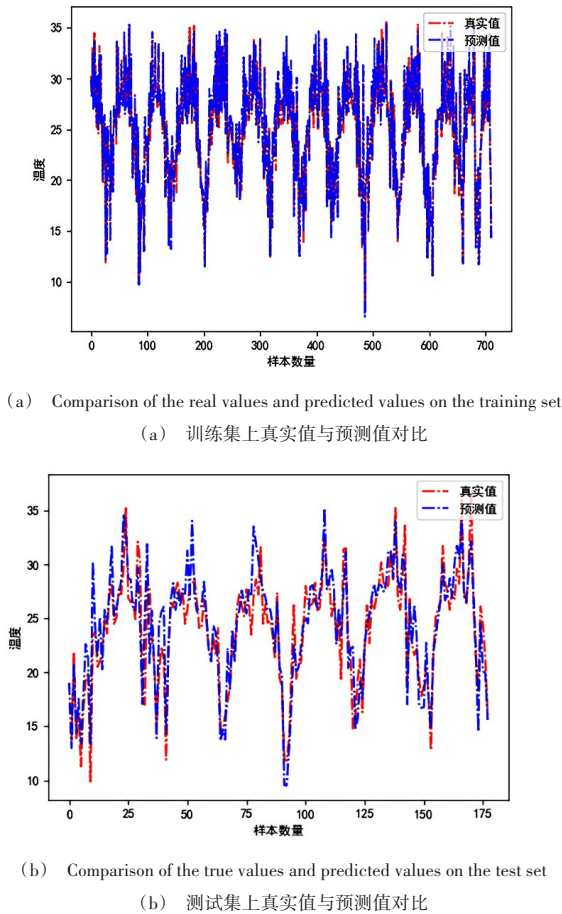


Fig. 7 Model prediction effect
图 7 模型预测效果

可以看出,CNN-GRU-ATT 模型在训练集和测试集上预测的气温变化曲线与实际气温曲线有着较高的一致性。这显示出模型能够较为准确地模拟实际气温波动,表明该模型在预测短期气温中展现出了较高的精确度。值得注意的是,模型在训练集上的表现略优于测试集,可能是因为训练集中包含了更多样本,从而使得模型能够更好地学习到气温变化规律。

3.4.3 Pearson 相关系数法对模型预测结果的影响

为了验证 Pearson 相关系数法在特征选择中的有效性,本文设计了消融实验,分别使用 CNN-GRU-ATT 模型和 CNN-GRU 模型进行测试。实验中,分别将通过 Pearson 相关系数法筛选的特征和未经过筛选的原始特征作为模型输入,以单步预测的方式进行未来 3 小时的气温预测。测试结果如表 2 所示。

可以看出,在某些情况下,减少输入特征的数量可以提高模型预测准确率。在数据预处理阶段,对两种模型结构(CNN-GRU 和 CNN-GRU-ATT)进行特征选择后, RMSE 和 MAE 都有所下降,表明进行 Pearson 相关系数法特征选择对 CNN-GRU-ATT 模型和 CNN-GRU 模型的预测精度均

Table 2 Pearson correlation coefficient method validity test results

表 2 Pearson 相关系数法有效性测试结果

是否特征选择	模型	RMSE	MAE
是	CNN-GRU	2.36	1.73
否	CNN-GRU	2.64	2.02
是	CNN-GRU-ATT	2.21	1.65
否	CNN-GRU-ATT	2.47	1.86

有很大提升,证明利用 Pearson 相关系数法进行特征选择有利于提高 CNN-GRU-ATT 模型和 CNN-GRU 模型预测精度。

3.4.4 不同时间步长对 CNN-GRU-ATT 模型预测效果的影响

为了探究时间步长对本文 CNN-GRU-ATT 模型预测效果的影响,开展了多步预测实验。多步预测是指在时间序列分析中,模型预测未来多个时间点的值,而不仅仅是下一个时间点的值。通过这一实验,旨在确定最佳预测时间步长,以提高气温预测准确性。预测误差结果如表 3 所示。

Table 3 Prediction errors at different time steps

表 3 不同时间步的预测误差

时间步长	RMSE	MAE
1	2.21	1.65
2	2.32	1.75
3	2.55	1.94
4	2.75	2.14
5	2.84	2.24

可以看出,模型在时间步长为 1 时的 RMSE 和 MAE 指标值最低,表明模型对最近时间点的预测最为准确,因此 CNN-GRU-ATT 模型最优步长为 1。随着时间的拉长,模型的 RMSE 和 MAE 指标值都逐步上升,表明预测性能也越差。这与气象预测规律一致,预测步长越长,能够提供的信息越少,使用模型学习到的信息也越少,这也表明本文模型可能更适合用于短期气温预测。

4 结语

本文针对传统气温预测方法在面对复杂的气候系统时预测精度不足这一问题,提出了一种基于 CNN-GRU-ATT 的气温预测模型,以开展海口市气温预测。该模型利用 Pearson 相关系数法对海口气象相关数据进行特征选择,再通过卷积神经网络(CNN)提取相关局部特征,捕捉多维历史数据中的有效信息。进一步地,采用门控循环单元(GRU)对特征序列进行时间序列建模,以识别气温变化的长期依赖性。最后,引入注意力机制(Attention)对 GRU 的输出进行加权,使模型能够集中关注重要的时间步,忽略不相关信息,从而优化预测效果。实验结果表明,本文算法在 RMSE 和 MAE 评价指标上均低于对比模型,证明了该模型的优越性,并为海口市气温短期预测提供了有效的解决方案。

在未来工作中,将尝试运用不同的特征选择方法进行特征选择,以进一步改进模型在中长期气温预测方面的性能。同时,通过增强算法对模型参数进行优化,并综合研究其对于组合模型的影响。

参考文献:

- [1] ZHAI Y H. Research on temperature prediction model based on BiLSTM and GCN[D]. Nanjing: Nanjing University of Posts and Telecommunications, 2022.
翟月昊. 基于BiLSTM和GCN的气温预测模型研究[D]. 南京:南京邮电大学, 2022.
- [2] HOU J, WANG Y, ZHOU J, et al. Prediction of hourly air temperature based on CNN-LSTM[J]. Geomatics, Natural Hazards and Risk, 2022, 13(1):1962-1986.
- [3] YANG H. Research on meteorological prediction based on deep learning[D]. Harbin: Harbin Institute of Technology, 2017.
杨函. 基于深度学习的气象预测研究[D]. 哈尔滨:哈尔滨工业大学, 2017.
- [4] YU T Y. Short-term temperature prediction model based on XGBoost and improved LSTM[D]. Wuhan: Huazhong University of Science and Technology, 2023.
于天滢. 基于XGBoost和改进LSTM的短期气温预测模型[D]. 武汉:华中科技大学, 2023.
- [5] DA Q, CHEN Y, DAI B, et al. Prediction of slope safety factor based on attention mechanism-enhanced CNN-GRU[J]. Sustainability, 2024, 16(15):1-23.
- [6] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural Computation, 1997, 9(8):1735-1780.
- [7] GERS F A, SCHMIDHUBER J, CUMMINS F. Learning to forget: continual prediction with LSTM[J]. Neural Computation, 2000, 12(10):2451-2471.
- [8] HUANG T, SHE Q, WANG Z, et al. GateNet: gating-enhanced deep network for click-through rate prediction [DB/OL]. <https://arxiv.org/abs/2007.03519>, 2020.
- [9] WANG Y J. Research on temperature prediction based on GCN-BiLSTM-attention fusion model[D]. Shijiazhuang: Hebei University of Economics and Business, 2024.
王玉洁. 基于GCN-BiLSTM-attention融合模型的气温预测研究[D]. 石家庄:河北经贸大学, 2024.
- [10] SAJJAD M, KHAN Z A, ULLAH A, et al. A novel CNN-GRU-based hybrid approach for short-term residential load forecasting[J]. IEEE Access, 2020, 8:143759-143768.
- [11] LIVIERIS I, PINTELAS E, PINTELAS A. A CNN-LSTM model for gold price time-series forecasting[J]. Neural Computing & Applications, 2020, 32:17351-17360.
- [12] LU W J, LI J Z, WANG J Y, et al. A novel model for stock closing price prediction using CNN-attention-GRU-attention[J]. Economic Computation & Economic Cybernetics Studies & Research, 2022, 56(3):251-264.
- [13] LI X X. Research on temperature prediction model in Guangxi based on neural network[D]. Nanning: Guangxi University for Nationalities, 2023.
李晓翔. 基于神经网络的广西气温预测模型研究[D]. 南宁:广西民族大学, 2023.
- [14] LIU H, YUE D M, MIAO G Y, et al. Short-term load forecasting technology for CNN-GRU hybrid neural network optimization based on Attention mechanism[J]. Electrotechnical Studies, 2024(9):20-23.
刘会, 岳东明, 苗光尧, 等. 基于Attention机制优化CNN-GRU混合神经网络的短期负荷预测技术[J]. 电工技术, 2024(9):20-23.
- [15] ZHAO P F, QIAN Y L, JIN X, et al. pH prediction of desulfurization system based on attention mechanism and CNN-GRU model[J]. Hydropower Energy Science, 2024, 42(9):199-203.
赵鹏飞, 钱玉良, 金鑫, 等. 基于注意力机制和CNN-GRU模型的脱硫系统pH值预测[J]. 水电能源科学, 2024, 42(9):199-203.
- [16] MA L Y, HU J C, DUAN X C, et al. Application of CNN-GRU model with attention mechanism in wind turbine abnormal state warning[J]. Journal of Nanjing University of Information Science & Technology, 2025, 17(3):374-383.
马良玉, 胡景琛, 段晓冲, 等. 具有注意力机制的CNN-GRU模型在风电机组异常状态预警中的应用[J]. 南京信息工程大学学报, 2025, 17(3):374-383.
- [17] MA S Z. Research on multi-series temperature prediction based on deep learning[D]. Lanzhou: Lanzhou University of Technology, 2022.
马司周. 基于深度学习的多序列气温预测研究[D]. 兰州:兰州理工大学, 2022.
- [18] CHEN L, ZHANG H L, WANG B, et al. Design of temperature prediction based on XGBoost and improved LSTNet[J]. Radio Engineering, 2023, 53(3):591-600.
陈岚, 张华琳, 汪波, 等. 基于XGBoost和改进LSTNet的气温预测设计[J]. 无线电工程, 2023, 53(3):591-600.
- [19] ZHOU W X, ZHAO L Y. Short-term traffic flow prediction based on attention mechanism and mixed model[J]. Journal of Chongqing University of Technology(Natural Science), 2025, 39(1):213-218.
周文学, 赵丽雅. 注意力机制与混合模型融合的短期交通流预测[J]. 重庆理工大学学报(自然科学), 2025, 39(1):213-218.
- [20] HU H, CHEN J P, LI Q, et al. Research on prediction of urban rainstorm water accumulation based on CNN-GRU-ATT[J]. Journal of North China University of Water Resources and Hydropower (Natural Science Edition), 2024, 45(4):27-35.
胡昊, 陈军朋, 李擎, 等. 基于CNN-GRU-ATT的城市暴雨积水预测研究[J]. 华北水利水电大学学报(自然科学版), 2024, 45(4):27-35.
- [21] WANG Y, YANG J B, LI R, et al. CNN-GRU temperature prediction model based on spatio-temporal series[J]. Journal of Computer Applications, 2023, 43(S2):54-59.
王益, 杨剑波, 李锐, 等. 基于时空序列的CNN-GRU气温预测模型[J]. 计算机应用, 2023, 43(S2):54-59.
- [22] JIN Y Q, LI Z H. Hourly temperature prediction based on LSTM neural network[J]. Modern Information Science and Technology, 2023, 7(9):26-29.
金郁淇, 李知何. 基于LSTM神经网络的逐时气温预报[J]. 现代信息科技, 2023, 7(9):26-29.

(责任编辑:孙娟)