

基于CNN-SAEDN-Res的短期电力负荷预测方法

崔杨¹,朱晗¹,王议坚¹,张璐²,李扬¹

(1. 东北电力大学 现代电力系统仿真控制与绿色电能新技术教育部重点实验室,吉林 吉林 132012;

2. 中国农业大学 信息与电气工程学院,北京 100083)

摘要:基于深度学习的序列模型难以处理混有非时序因素的负荷数据,这导致预测精度不足。提出一种基于卷积神经网络(CNN)、自注意力编码解码网络(SAEDN)和残差优化(Res)的短期电力负荷预测方法。特征提取模块由二维卷积神经网络组成,用于挖掘数据间的局部相关性,获取高维特征。初始负荷预测模块由自注意力编码解码网络和前馈神经网络构成,利用自注意力机制对高维特征进行自注意力编码,获取数据间的全局相关性,从而模型能根据数据间的耦合关系保留混有非时序因素数据中的重要信息,通过解码模块进行自注意力解码,并利用前馈神经网络回归初始负荷。引入残差机制构建负荷优化模块,生成负荷残差,优化初始负荷。算例结果表明,所提方法在预测精度和预测稳定性方面具有优势。

关键词:短期电力负荷预测;卷积神经网络;自注意力机制;残差机制;负荷优化

中图分类号:TM73

文献标志码:A

DOI:10.16081/j.epae.202308018

0 引言

提高负荷预测的准确度和可靠度有利于电网经济、安全运行。目前,短期负荷预测方法主要分为两大类,即基于时序分析的统计学方法和基于数据驱动的机器学习方法。统计学方法于20世纪70年代被引入,该类方法在不考虑相关因素对负荷影响的前提下,通过分析历史负荷数据探索负荷的周期性变化^[1-2],虽然能有效处理平稳的负荷时间序列变化,但无法有效处理天气因素骤变、设备故障检修、特殊活动等引起的非线性负荷变化,即非线性学习能力差。随着智能计算的发展,基于数据驱动的机器学习方法被应用于负荷预测领域^[3],该类方法通过对包含负荷、气象、日期等在内的各种数据进行处理,提取并选择能够充分反映未来负荷变化的特征,使负荷变化与外界因素相关联,从而实现高精度预测。

当前,应用于负荷预测的机器学习方法主要包括传统的机器学习方法^[4-5]和深度学习模型。传统的机器学习方法需要先开展特征工程工作,对数据进行深度分析后才能建立模型进行学习,否则预测效果会不理想。相比之下,深度学习模型可以直接

将数据传递到网络,利用其强大的非线性表达能力挖掘特征,建立负荷与各类因素之间的关联性,实现高精度预测。

在深度学习模型中,最具代表性的是卷积神经网络(convolutional neural network, CNN)和序列模型(循环神经网络及其变体)^[6-7],不少研究者基于此提出负荷预测模型,利用长短期记忆(long short-term memory, LSTM)模型^[8]、门控循环单元(gate recurrent unit, GRU)模型^[9]等挖掘负荷数据的时序性特点。考虑到序列模型不能挖掘特征在高维空间的联系,部分研究者将CNN和序列模型相结合组成混合神经网络,由CNN提取高维特征,并将其交由序列模型处理^[10-13]。需要注意的是,序列模型仅适合处理时间序列数据,若输入数据中混入日期、气象等非时序因素,则模型将无法发挥其特性,不能通过挑选重要信息使之前的输入对未来的输出产生正确的影响,即无法类似时间序列建立前后信息之间的耦合关系,不能对预测结果产生积极的推动作用。

为此,本文提出一种基于CNN、自注意力^[14]编码解码网络(self-attention encoder-decoder network, SAEDN)以及残差优化(residual-refinement, Res)的短期电力负荷预测方法。该方法可不受数据类型的限制,利用自注意力机制建立特征中任意信息之间的耦合关系,并根据耦合程度自适应地选择保留重要特征或遗忘不重要特征。此外,引入残差机制,在负荷回归之后通过添加负荷优化模块优化初始负荷,进一步提升方法预测精度。

1 基于CNN-SAEDN-Res的负荷预测方法

1.1 模型概述

本文所提负荷预测方法中的模型整体结构如图

收稿日期:2022-12-21;修回日期:2023-06-07

在线出版日期:2023-08-15

基金项目:现代电力系统仿真控制与绿色电能新技术教育部重点实验室开放课题(MPSS2021-09);吉林省自然科学基金资助项目(YDZJ202101ZYTS149)

Project supported by the Open Fund of the Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education(MPSS2021-09) and the Natural Science Foundation of Jilin Province (YDZJ202101ZYTS149)

1所示,包含特征提取模块、初始负荷预测模块和负荷优化模块。预测流程为:首先,通过特征提取模块提取特征,构建数据之间的邻近(局部)特征关系;其次,对提取的特征进行二维空间位置编码,对其赋予空间位置信息;然后,将编码后的特征送入SAEDN中进行编码和解码,再通过前馈神经网络(feedforward neural network, FFN)回归日负荷数据;最后,将初始预测负荷与原始数据一同送入负荷优化模块中,以优化初始预测负荷。

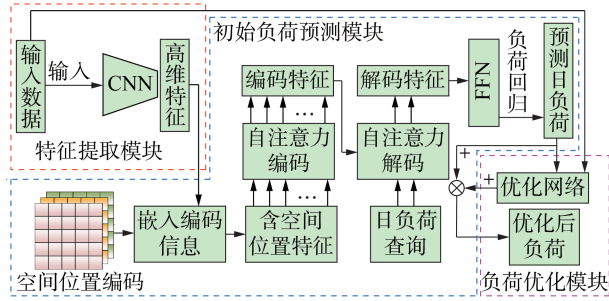


图1 负荷预测模型的整体结构

Fig.1 Overall architecture of load forecasting model

1.2 特征提取模块

在负荷预测中,CNN常用于挖掘数据之间的局部相关性,获取高维特征向量。然而,利用现有方法^[10-13]设计的CNN通常很浅,难以获取较高级的语义信息。为此,本文设计一种较深的CNN,将其作为特征提取模块,以挖掘语义信息,提高预测精度,具体结构如附录A图A1所示。

特征提取模块共包含7个卷积层,卷积核大小为3×3,在每层卷积后均跟随1个批归一化层和1个激活层。除了第2个卷积层采用步长为2的卷积操作外,其余卷积层的步长均为1。此外,在模型中引入残差连接,以防止梯度消失并加快模型收敛^[15]。

1.3 初始负荷预测模块

初始负荷预测模块是本文模型的核心部分,其性能直接影响模型的预测精度。目前的不少预测方法使用序列模型处理数据以及预测负荷,然而,在混杂了非序列因素的数据中,序列模型难以有效地建立前后信息之间的联系,导致预测精度下降。为了避免该问题,本文在初始负荷预测模块中舍弃了序列模型,采用SAEDN对特征进行编码和解码,并通过FFN^[16]实现负荷回归。初始负荷预测模块如附录A图A2所示,由空间位置编码、多个自注意力编码解码和FFN组成。

1.3.1 空间位置编码

序列模型的独特结构使其能够捕捉数据之间的序列关系,但SAEDN无法自主获取数据的位置信息,这使得特定时间与具有特定属性的负荷数据(负荷峰谷)之间难以关联,因此,需要进行位置编码,以

确保数据位置的严格对应。具体公式为:

$$\begin{cases} V_{PE}(p, 2r) = \sin \frac{p}{10\,000^{2r/d_{model}}} \\ V_{PE}(p, 2r+1) = \cos \frac{p}{10\,000^{2r/d_{model}}} \end{cases} \quad (1)$$

式中: $V_{PE}(p, 2r)$ 为位置编码矩阵 V_{PE} 的第 p 行第 $2r$ 列元素, p 为数据在序列中的位置, r 为向量维度的索引; d_{model} 为特征最大长度。

由于特征提取模块的输出为二维特征而非序列结构,因此需要分别对2个方向进行1次位置编码。在编码完成后,将位置信息嵌入输出特征中,以生成包含空间位置信息的高维特征,可将这些特征作为自注意力编码模块的输入。

1.3.2 自注意力机制

在电力负荷预测中,注意力机制被用于关注重要信息和特征^[17-18],但注意力机制仅关注因素信息与预测负荷之间的关系,其作用是局部的,忽视了各因素内部、各因素信息之间的关键信息,这不利于包含时间信息的负荷预测。负荷变化通常具有一定的平滑属性,其趋势可以通过邻近负荷反映,因此,为了建立数据之间的相关性,需要构造自注意力模型,并通过多次并行运行构造多个自注意力,形成多头自注意力层。自注意力机制的具体流程如附录A图A3所示。

首先,利用特征向量 X 与矩阵系数 W^q 、 W^k 、 W^v 分别计算得到矩阵 Q 、 K 、 V ,具体计算公式为:

$$\begin{cases} Q = W^q X \\ K = W^k X \\ V = W^v X \end{cases} \quad (2)$$

然后,将 K 转置后与 Q 相乘,计算得到特征相关性,再通过Softmax函数对特征相关性进行归一化处理,得到注意力权重 α 。若 $K^T Q$ 的结果过大,则会导致Softmax函数处理后的偏导数过小,使梯度趋近于0,不利于模型收敛,为此,在归一化处理之前改变尺度。上述计算公式为:

$$\alpha = \text{Softmax} \left(\frac{K^T Q}{\sqrt{d_k}} \right) \quad (3)$$

式中: d_k 为矩阵 K 的列数。

最后,根据 V 与 α 计算得到自注意力 A ,即:

$$A = V\alpha \quad (4)$$

由上述计算过程可知,自注意力机制通过输入向量与自身转置相乘,获取输入向量在自身上的投影。投影值越大,表明数据的关联程度越高。自注意力机制通过这种方式可以不受数据类型的限制,即使数据中混入了非序列因素,依然可以通过全局信息的作用关系选择重要的信息特征,遗忘不重要的部分。

1.4 负荷优化模块

残差是指实际值与理论计算值之间的绝对误差,其产生原因多种多样,如星期类型的改变、气象因素的骤变等均可能导致模型预测不准,产生较大的残差,但本质原因在于模型的学习能力不足,难以有效拟合日负荷曲线。针对该问题,本文在初始负荷预测结束后通过添加负荷优化模块生成残差,优化初始预测负荷的残差分布,从而改善第2次的负荷预测结果。该方法的数学原理如下。

假设 y_{init} 、 y_{Refine} 、 e^* 分别为初始负荷预测结果、第2次负荷预测结果、负荷优化模块生成的残差,三者之间应满足:

$$e^* = y_{\text{Refine}} - y_{\text{init}} \quad (5)$$

预测值 y 、实际值 $\hat{y}$ 、残差 e 之间满足:

$$y = \hat{y} - e \quad (6)$$

联立式(5)和式(6),可得:

$$e^* = e_{\text{init}} - e_{\text{Refine}} \quad (7)$$

式中: e_{init} 、 e_{Refine} 分别为初始负荷预测误差和第2次负荷预测误差。

式(7)表明,负荷优化模块学习到的是一个残差值,旨在优化第2次负荷预测的残差结果。

本文所设计的负荷优化模块如附录A图A4所示。与特征提取模块不同的是,在负荷优化模块中,不是将输入特征当作二维结构利用卷积进行处理,而是利用特征的序列结构,采用GRU紧跟一组FFN,生成负荷残差,优化初始负荷,这样处理是由于若在优化模块中再采用CNN-SAEDN结构,则会大幅增加模型的时间和空间复杂度,降低优化模块的性价比。

2 仿真设置

2.1 仿真流程

本文所提短期负荷预测方法的流程如附录A图A5所示。首先,进行数据预处理,构建负荷预测输入特征集。然后,将特征集划分为训练样本和测试样本,将训练样本送入预测模型中,设置训练超参数并开始训练,根据损失函数计算训练损失。在每轮训练结束后,均要判断训练轮数是否达到预设值:若未达到预设值,则继续训练,并用优化器算法更新模型权重;若达到预设值,则停止训练,将测试样本送入预测模型中,计算模型性能并预测负荷。

2.2 数据预处理

2.2.1 异常数据处理

电力负荷数据会受到人为失误、设备故障等多种因素的影响,从而导致数据出现异常值,这些异常值会影响模型的处理,进而影响预测效果,为此,本文对这些异常数据进行筛选、删除、替换处理,具体方法如下。

假设某日负荷数据为 $\{x_1, x_2, \dots, x_h\}$, x_a ($a=1, 2, \dots, h$) 为第 a 次采集的负荷数据, $h=24$ 。

由于异常数据属于离群点,会加大数据整体的离散程度,而标准差可以反映该特性,因此,本文通过计算标准差判断数据集中是否存在异常数据。标准差 σ 的计算公式为:

$$\sigma = \sqrt{\frac{1}{h} \sum_{a=1}^h (x_a - \mu)^2} \quad (8)$$

$$\mu = \frac{1}{h} \sum_{a=1}^h x_a \quad (9)$$

式中: μ 为样本均值。

若 $\sigma > 140$ (该数据通过10组异常数据和随机抽取的100组正常数据观察得出),则认为该日负荷数据集离散度过高,存在异常数据,但此时依然无法定位异常数据的具体位置。

由于日负荷数据不服从正态分布,因此需要手工设定阈值,通过数据远离平均距离的标准差倍数判断该数据是否为异常数据。数据 x_a 远离平均距离的标准差比例 m_a 的计算公式为:

$$m_a = \frac{|x_a - \mu|}{\sigma} \quad (10)$$

若 $m_a > 1$,则表明数据 x_a 为异常数据,求取异常数据 x_a 左右正常数据的均值对 x_a 进行填充替换。

2.2.2 数据归一化

数据归一化将量纲不同且数值差异较大的数据映射到 $[0, 1]$,能够加快模型收敛,减小预测误差。归一化计算公式为:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (11)$$

式中: x^* 为归一化后的数据; x 为归一化前的数据; $x_{\max}$ 、 $x_{\min}$ 分别为一组数据中的最大值和最小值。

2.2.3 负荷预测输入特征集的构建

合理数据结构的构建有助于模型建立数据之间的关系,由于电力负荷受到多重因素的影响,因此需要综合考虑这些因素,构建多维特征向量并将其作为模型输入特征,具体如附录A表A1所示。

在各影响因素中,将日期因素(除时间外)采用独热编码重新表示,便于模型进行数据处理,最终特征向量的空间维度为 9×24 ,其中,由于日期因素的范围为1~31日,用2行向量表示其独热编码。3.1节算例分析所用数据集中某一日的输入特征如附录A图A6所示。

2.3 仿真环境与超参数设置

本文所提负荷预测模型分为两部分训练:第1部分为基本预测网络CNN-SAEDN,使用Adam优化器,批样本数设置为32,初始学习率设置为0.001,在第150轮和第300轮时学习率较之前均减半,共训练500轮;第2部分为负荷优化网络,使用Adam优化

器,批样本数设置为16,初始学习率设置为0.01,在第100轮和第200轮时学习率较之前均减半,共训练300轮。为了证明本文所提负荷预测方法的有效性,所有模型均基于网格搜索法调整超参数,以使模型性能均能达到近似最优。此外,所有仿真均在单张GTX850M上进行,编程语言为Python3.7,深度学习框架为Pytorch,保证模型预测精度和速度对比的公平性。

2.4 损失函数

本文采用具有易求导、存在解析解等优点的均方损失函数计算损失,函数表达式为:

$$F_{\text{loss}} = \frac{1}{b} \sum_{j=1}^b F_j \quad (12)$$

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

式中: F_{loss} 为批样本数 b 下的总训练损失; F_j 为第 j 个样本下的日负荷均方损失; n 为日负荷实际考核点数; $\hat{y}_i$ 为考核点 i 的实际负荷; y_i 为考核点 i 的预测负荷。

计算日负荷损失是由于本文模型采用的预测策略是多对多预测,即通过1次预测直接输出次日24 h的负荷预测结果。

2.5 性能指标

依据电力行业负荷预测技术规范^[19],选择日平均负荷预测准确率 A_d 反映模型预测精度,日平均误差 E_d 反映预测的整体性偏差,计算公式分别为:

$$A_d = \left[1 - \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{\hat{y}_i} \right)^2} \right] \times 100 \% \quad (14)$$

$$E_d = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (15)$$

3 算例分析

为验证本文所提方法的科学性和有效性,利用数据集对所提方法中的模型进行仿真,并与其他方法中的模型进行比较与分析。

3.1 美国公共事业公司数据集算例

本文采用美国公共事业公司(Global Energy Forecasting Competition, GEFCom)数据集2004年1月1日至2008年6月29日的每小时负荷数据进行仿真^[20],该数据集中除了包含负荷信息外,还包含日期、时间、节假日和气温信息。

本文将数据集中2004年1月1日至2008年6月22日的数据组成训练样本集(39 216条样本),对2008年6月23日至2008年6月29日的负荷进行日前预测(提前1 d直接预测次日24 h的负荷),并利用2.5节中的指标对模型进行量化分析。为了使对比更加全面,本文将预测结果分为4个部分进行比较,

分别为1周中的第1天预测结果、5 d工作日的平均预测结果、2 d休息日的平均预测结果和1周的平均预测结果。

3.1.1 模型消融仿真结果

本文模型主要由特征提取模块、初始负荷预测模块和负荷优化模块组成,因此,在仿真中,通过消融模块分别展现各部分对预测精度的提升效果。本文模型在GEFCom数据集上的消融仿真结果如表1所示。相较于模型1,添加初始负荷预测模块后,模型2的预测准确率提高了0.91个百分点,平均误差降低了2.92 kW,这表明在该结构下模型预测结果更加准确和平滑。

表1 本文模型在GEFCom数据集上的消融仿真结果

Table 1 Ablation simulative results of proposed model on GEFCom dataset

模型	模块设置			$A_d / \%$	E_d / kW
	特征提取模块	初始负荷预测模块	负荷优化模块		
1	√			95.36	17.17
2	√	√		96.27	14.25
3	√	√	√	96.37	13.72

注:“√”表示模型中包含相应部分。

为了验证负荷优化模块的作用,本文对2次预测负荷的残差结果进行正态性检验^[21]。预测结果优化前后的标准化残差分布如图2所示。由图可知,红色的点有11个落在标准差范围之外,而蓝色的点仅有8个落在标准差范围之外。根据正态分布的理论,仿真点的标准化残差落在标准差范围之外的概率应小于5%,即在本仿真中异常点的数量应少于8.4(168×5%)个。初始负荷的预测结果没有通过残差的正态性检验,模型对日负荷曲线的拟合能力不足,而优化后的结果则满足该检验要求。结合表1可知,添加负荷优化模块后模型预测准确率提高了0.1个百分点,平均误差降低了0.53 kW。上述结果表明,负荷优化模块优化了负荷的残差分布,改善了模型的拟合结果,提高了预测精度。

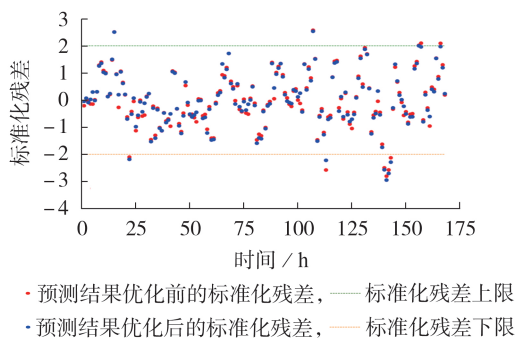


图2 预测结果优化前后的标准化残差分布

Fig.2 Distribution of standardized residuals before and after optimization of forecasting results

3.1.2 不同模型的对比仿真

3.1.2.1 仿真说明

本节设置2组仿真,用于验证本文模型性能:第1组仿真是本文模型与LSTM、GRU、CNN-LSTM、CNN-GRU的对比,主要用于验证本文模型相较于序列模型具有更优的预测效果;第2组仿真是本文模型与空间注意力(spatial attention, SA)机制^[22]、通道注意力(channel attention, CA)机制^[23]的对比,主要用于验证自注意力机制能取得更优的预测结果。

3.1.2.2 与序列模型的对比仿真

在该仿真中,所有模型使用的CNN与CNN-SAEDN中的结构相同,将LSTM和GRU均设置为3层,通过相同结构的FFN得到日负荷预测结果。LSTM、CNN-LSTM、GRU、CNN-GRU的批样本数均设置为16,初始学习率为0.01,训练轮数和递减效率与CNN-SAEDN一致。

仿真结果如附录A表A2所示。由表可知:相较于CNN、LSTM、CNN-LSTM、GRU、CNN-GRU,在工作日,本文模型的预测准确率分别提高了0.93、1.23、1.11、1.00、0.72个百分点,平均误差分别降低了3.28、5.43、4.11、3.12、3.20 kW;在休息日,本文模型的预测准确率和平均误差依然取得了最优的预测结果;综合一周的平均值进行分析,本文模型的预测准确率和平均误差优势明显,这表明本文模型的预测精度和稳定性优于对比模型。

各模型在周一的负荷预测结果如图3所示。由图可知,除了本文模型之外,其他模型预测的负荷高峰时间均向后偏移,这说明通过自注意力机制有效挖掘了负荷高峰与气温、时间之间的联系,克服了因序列模型无法有效处理非时序因素而导致的预测滞后问题。

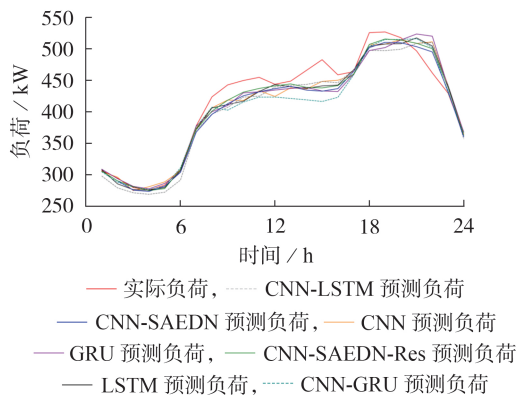


图3 各模型在周一的负荷预测结果

Fig.3 Load forecasting results of each model on Monday

CNN-SAEDN和序列模型的周负荷曲线图如附录A图A7—A11所示。由图可知,序列模型的结果易出现滞后现象,尤其是在负荷高峰较为单一的工

作日。

3.1.2.3 与注意力机制的对比仿真

在该仿真中,所有模型均是在3.1.2.2节CNN、CNN-LSTM、CNN-GRU的基础上添加空间注意力和通道注意力,将注意力模块添加到CNN的残差模块中,除了初始学习率为0.001外,模型训练的超参数与3.1.2.2节一致。

仿真结果见附录A表A3。在CNN-SAEDN-Res中剔除优化模块,仅对CNN-SAEDN的预测结果与注意力机制下的各模型结果进行比较,突出自注意力机制相比其他注意力机制的优势。在含有空间注意力和通道注意力的模型中,根据1周的预测结果可知,表现最好的是基于CNN-LSTM的空间注意力(CNN-LSTM-SA)和基于CNN-GRU的通道注意力(CNN-GRU-CA)。相较于CNN-LSTM-SA和CNN-GRU-CA:在工作日,本文模型的预测准确率分别提高了0.55、0.73个百分点;综合分析1周的预测结果,本文模型的预测准确率分别提高了0.47、0.57个百分点,平均误差也有所降低,本文模型性能明显更优,这表明在负荷预测中数据间的全局注意力所构造的全局相关性比数据与预测结果间的局部相关性更能挖掘重要的数据特征,建立与预测结果之间的非线性关系。

3.2 我国南方某地区算例

本算例采用我国南方某地区数据集,该数据集包含2016年1月1日至2018年12月31日的电力负荷数据,一天采样24次,采样间隔为1 h。此外,该数据集还包含各日的最高气温、最低气温、平均气温、相对湿度和降雨量。本文将2016年1月1日至2018年12月24日的数据组成训练样本集(26 112条样本),对2018年12月25日至2018年12月31日的负荷进行日前预测。为了体现模型的泛化性能,模型参数设置与3.1节相同。

各模型在该数据集上1周的平均预测准确率和平均误差以及各模型进行单次日预测所消耗的时间如附录A表A4所示。由表可知,本文模型的预测准确率最高且平均误差最小,但本文模型的计算速度远慢于其他模型,这是由于在SAEDN中进行了多次并行运算,从而导致模型的参数过多以及计算量过大。但本文模型的运算时长依然控制在毫秒级别,进行单次日负荷预测仅需19.1 ms,满足负荷预测的实时性需求。

4 结论

针对因序列模型不能有效处理含有非时序因素的负荷数据而导致预测精度不足的问题,本文提出一种基于CNN-SAEDN-Res的短期日负荷预测方法。通过算例验证得出如下结论。

1)自注意力机制可以不受数据类型限制,即使是包含非时序因素的数据,也能通过全局注意力建立各数据之间的联系,为数据动态分配权重,自适应地保留或遗忘信息,在非时序数据中也能起到序列模型在时序数据中类似的效果。

2)对残差的分析和正态性检验结果说明可以在残差层面对负荷预测结果进行优化,而不仅限于通过加强模型的设计提升模型的非线性学习能力来提高预测质量。

3)残差机制不仅体现在残差处理上,还体现在负荷预测结果的多阶段优化。算例结果表明,通过增加优化步骤可以提高模型的预测精度和稳定性。

本文方法尚存在不足之处,计算速度的对比结果表明,本文方法中的模型参数过多,未来笔者将考虑通过剪枝技术进行结构优化。

附录见本刊网络版(<http://www.epae.cn>)。

参考文献:

- [1] 谢毓广,刘童,陈凡,等. 基于置信度理论和多元线性回归方程的电网负荷预测[J]. 电工技术,2018(13):42-45.
- XIE Yuguang, LIU Tong, CHEN Fan, et al. Network load forecasting based on confidence theory and multiple linear regression equation[J]. Electric Engineering, 2018(13):42-45.
- [2] 赵峰,孙波,张承慧. 基于多变量相空间重构和卡尔曼滤波的冷热电联供系统负荷预测方法[J]. 中国电机工程学报,2016,36(2):399-406.
- ZHAO Feng, SUN Bo, ZHANG Chenghui. Cooling, heating and electrical load forecasting method for CCHP system based on multivariate phase space reconstruction and Kalman filter[J]. Proceedings of the CSEE, 2016, 36(2):399-406.
- [3] 范士雄,李立新,王松岩,等. 人工智能技术在电网调控中的应用研究[J]. 电网技术,2020,44(2):401-411.
- FAN Shixiong, LI Lixin, WANG Songyan, et al. Application analysis and exploration of artificial intelligence technology in power grid dispatch and control[J]. Power System Technology, 2020, 44(2):401-411.
- [4] 孙超,吕奇,朱思瞳,等. 基于双层XGBoost算法考虑多特征影响的超短期电力负荷预测[J]. 高电压技术,2021,47(8):2885-2898.
- SUN Chao, LÜ Qi, ZHU Sitong, et al. Ultra-short-term power load forecasting based on two-layer XGBoost algorithm considering the influence of multiple features[J]. High Voltage Engineering, 2021, 47(8):2885-2898.
- [5] 吴倩红,高军,侯广松,等. 实现影响因素多源异构融合的短期负荷预测支持向量机算法[J]. 电力系统自动化,2016,40(15):67-72,92.
- WU Qianhong, GAO Jun, HOU Guangsong, et al. Short-term load forecasting support vector machine algorithm based on multi-source heterogeneous fusion of load factors[J]. Automation of Electric Power Systems, 2016, 40(15):67-72, 92.
- [6] 肖白,高文瑞,李道明,等. 基于3 σ -CEEMDAN-LSTM的空间负荷预测方法[J]. 电力自动化设备,2023,43(3):159-165.
- XIAO Bai, GAO Wenrui, LI Daoming, et al. Spatial load forecasting method based on 3 σ -CEEMDAN-LSTM[J]. Electric Power Automation Equipment, 2023, 43(3):159-165.
- [7] 郑瑞晓,张姝,肖先勇,等. 考虑温度模糊化的多层长短时记忆神经网络短期负荷预测[J]. 电力自动化设备,2020,40(10):181-186.
- ZHENG Ruixiao, ZHANG Shu, XIAO Xianyong, et al. Short-term load forecasting of multi-layer long short-term memory neural network considering temperature fuzziness[J]. Electric Power Automation Equipment, 2020, 40(10):181-186.
- [8] 孙辉,杨帆,高正男,等. 考虑特征重要性值波动的MI-BiLSTM短期负荷预测[J]. 电力系统自动化,2022,46(8):95-103.
- SUN Hui, YANG Fan, GAO Zhengnan, et al. Short-term load forecasting based on mutual information and bi-directional long short-term memory network considering fluctuation in importance values of features[J]. Automation of Electric Power Systems, 2022, 46(8):95-103.
- [9] 王增平,赵兵,纪维佳,等. 基于GRU-NN的短期负荷预测方法[J]. 电力系统自动化,2019,43(5):53-58.
- WANG Zengping, ZHAO Bing, JI Weijia, et al. Short-term load forecasting method based on GRU-NN model[J]. Automation of Electric Power Systems, 2019, 43(5):53-58.
- [10] 姚程文,杨苹,刘泽健. 基于CNN-GRU混合神经网络的负荷预测方法[J]. 电网技术,2020,44(9):3416-3424.
- YAO Chengwen, YANG Ping, LIU Zejian. Load forecasting method based on CNN-GRU hybrid neural network[J]. Power System Technology, 2020, 44(9):3416-3424.
- [11] 朱凌建,荀子涵,王裕鑫,等. 基于CNN-BiLSTM的短期电力负荷预测[J]. 电网技术,2021,45(11):4532-4539.
- ZHU Lingjian, XUN Zihan, WANG Yuxin, et al. Short-term power load forecasting based on CNN-BiLSTM[J]. Power System Technology, 2021, 45(11):4532-4539.
- [12] 邹智,吴铁洲,张晓星,等. 基于贝叶斯优化CNN-BiGRU混合神经网络的短期负荷预测[J]. 高电压技术,2022,48(10):3935-3945.
- ZOU Zhi, WU Tiezhou, ZHANG Xiaoxing, et al. Short-term load forecast based on Bayesian optimized CNN-BiGRU hybrid neural networks[J]. High Voltage Engineering, 2022, 48(10):3935-3945.
- [13] 陆继翔,张琪培,杨志宏,等. 基于CNN-LSTM混合神经网络模型的短期负荷预测方法[J]. 电力系统自动化,2019,43(8):131-137.
- LU Jixiang, ZHANG Qipei, YANG Zhihong, et al. Short-term load forecasting method based on CNN-LSTM hybrid neural network model[J]. Automation of Electric Power Systems, 2019, 43(8):131-137.
- [14] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C]//31st Conference on Neural Information Processing Systems. Long Beach, USA:MIT Press, 2017:6000-6010.
- [15] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Las Vegas, USA:IEEE, 2016:770-778.
- [16] 蔡秋娜,潮铸,苏炳洪,等. 基于一种新型鲁棒损失的神经网络短期负荷预测方法[J]. 电网技术,2020,44(11):4132-4139.
- CAI Qiuna, CHAO Zhu, SU Binghong, et al. Short-term load forecasting method based on a novel robust loss neural network algorithm[J]. Power System Technology, 2020, 44(11):4132-4139.
- [17] 王继东,杜冲. 基于Attention-BiLSTM神经网络和气象数据修正的短期负荷预测模型[J]. 电力自动化设备,2022,42(4):172-177,224.
- WANG Jidong, DU Chong. Short-term load prediction model based on Attention-BiLSTM neural network and meteorological data correction[J]. Electric Power Automation Equipment, 2022, 42(4):172-177, 224.
- [18] 姚程文. 基于深度学习的短期电力负荷预测方法研究[D]. 广州:华南理工大学,2021.

- YAO Chengwen. Research on short-term power load forecasting methods based on deep learning[D]. Guangzhou: South China University of Technology, 2021.
- [19] 国家能源局. 电网短期和超短期负荷预测技术规范: DL/T 1711—2017[S]. 北京: 中国电力出版社, 2017.
- [20] BEN T S, HYNDMAN R J. A gradient boosting approach to the Kaggle load forecasting competition[J]. International Journal of Forecasting, 2014, 30(2): 382-394.
- [21] 胡菊华. 基于残差分析的线性回归模型的诊断与修正[J]. 统计与决策, 2019, 35(24): 5-8.
- HU Juhua. Diagnosis and correction of linear regression model based on residual analysis[J]. Statistics & Decision, 2019, 35(24): 5-8.
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[M] // Computer Vision-ECCV 2018. Cham, Switzerland: Springer International Publishing, 2018: 3-19.
- [23] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.

作者简介:

崔 杨(1980—), 男, 教授, 博士研究生导师, 博士, 主要研究方向为电力系统运行分析、新能源联网发电关键技术等 (E-mail: cuiyang0432@163.com)。

(编辑 王锦秀)

Short-term power load forecasting method based on CNN-SAEDN-Res

CUI Yang¹, ZHU Han¹, WANG Yijian¹, ZHANG Lu², LI Yang¹

(1. Key Laboratory of Modern Power System Simulation and Control & Renewable Energy Technology, Ministry of Education, Northeast Electric Power University, Jilin 132012, China;

2. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China)

Abstract: The sequence model based on deep learning is difficult to deal with load data mixed with non-temporal factors, which leads to insufficient forecasting precision. A short-term power load forecasting method based on convolutional neural network (CNN), self-attention encoder-decoder network (SAEDN) and residual-refinement (Res) is proposed. The feature extraction module is composed of a two-dimensional convolutional neural network, which is used to mine the local correlation between data and obtain the high-dimensional features. The initial load forecasting module is composed of a self-attention encoder-decoder network and a feedforward neural network. The self-attention mechanism is used for self-attention encoding of high-dimensional features, the global correlation between data is obtained, thus the model is able to retain important information in the data mixed with non-temporal factors according to the coupling relationship between data. The self-attention decoding is performed by the decoding module, and the feedforward neural network is used for the initial load regression. The residual mechanism is introduced to build the load optimization module, the residual load is generated, and the initial load is optimized. The example results show that the proposed method has advantages in terms of forecasting accuracy and forecasting stability.

Key words: short-term power load forecasting; convolutional neural network; self-attention mechanism; residual mechanism; load optimization

附录 A:

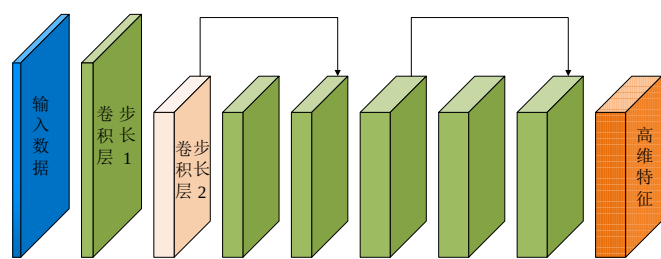


图 A1 特征提取模块结构
Fig.A1 Structure of feature extraction module

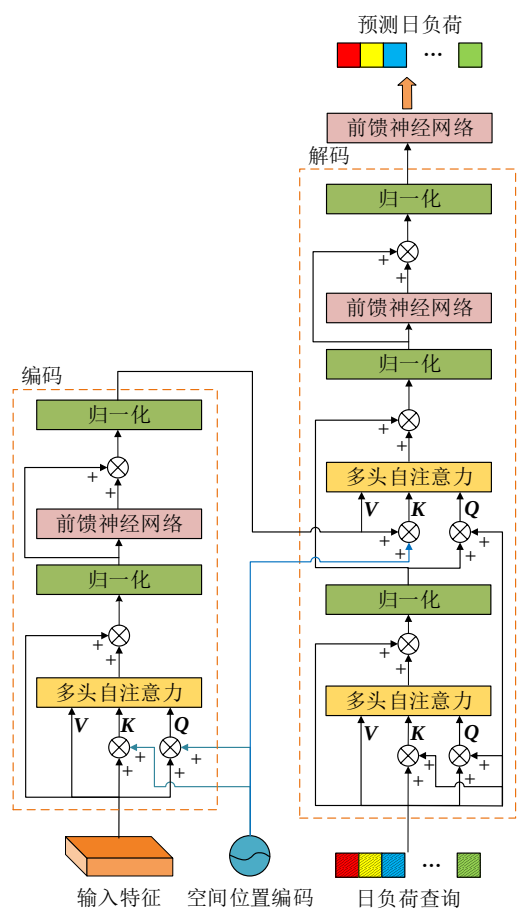


图 A2 初始负荷预测模块结构图
Fig.A2 Structure of initial load forecasting module

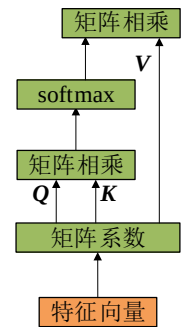


图 A3 自注意力机制
Fig.A3 Self-attention mechanism

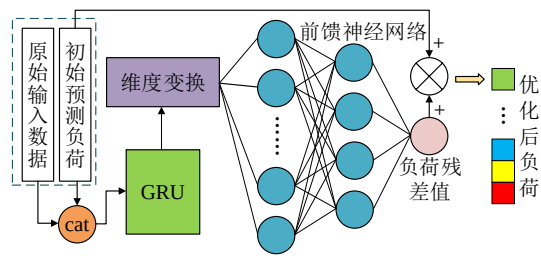


图 A4 负荷优化模块结构
Fig.A4 Structure of load refinement module

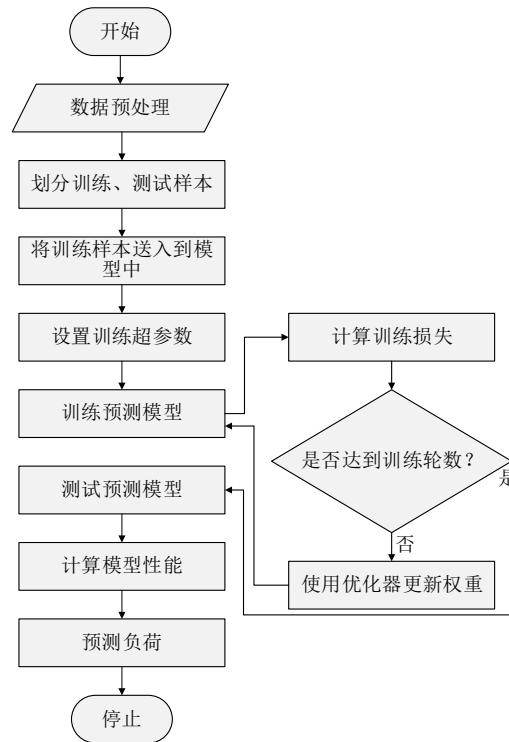


图 A5 短期负荷预测方法流程图
Fig.A5 Flowchart of short-term load forecasting method

表 A1 输入特征集
Table A1 Input feature set

影响因素	因素类型	描述
日期因素	月份	1~12 月
	日期	1~31 日
	星期	星期一至星期日
	小时	1~24h
	工作日/休息日	休息日包括周末和节假日
气象因素	气温/湿度/降雨量	24 h 气温/日最高、最低、平均气温/平均湿度、降雨量
相似负荷因素	前一个工作日/休息日负荷	根据数据集不同单位为 kW/MW
	前二个工作日/休息日负荷	

月份	1	0	0	0	0	0	0	0	0
日期	0	0	1	0	0	0	0	0	0
	0	0	0	0	0	0	0	0	0
星期	0	0	0	0	0	1	0	0	0
休息日	0	1	0	0	0	0	0	0	0
小时	1	2	3	4	5	6	7	23	24
温度	48	48	47	49	49	49	49	62	62
前一相似日负荷	461	438	447	439	448	485	534	553	481
前二相似日负荷	484	457	450	448	444	490	523	550	522

图 A6 模型输入特征
Fig.A6 Input feature of model

表 A2 本文模型与序列模型的预测结果比较

Table A2 The prediction results of the model in this paper are compared with those of the sequence model

方法	第 1 天		工作日		休息日		一周	
	A_d /%	E_d /kW	A_d /%	E_d /kW	A_d /%	E_d /kW	A_d /%	E_d /kW
CNN-SAEDN-Res	96.28	12.32	96.58	12.30	95.82	17.30	96.37	13.72
CNN-SAEDN	95.79	14.42	96.49	12.89	95.73	17.65	96.27	14.25
CNN	95.89	14.27	95.65	15.58	94.64	21.16	95.36	17.17
LSTM	95.75	14.88	95.35	17.73	95.01	20.47	95.25	18.51
CNN-LSTM	95.14	17.98	95.47	16.41	95.57	17.55	95.50	16.74
GRU	95.14	16.54	95.58	15.42	94.79	21.05	95.35	17.03
CNN-GRU	94.48	18.28	95.86	15.50	94.46	21.41	95.46	17.19

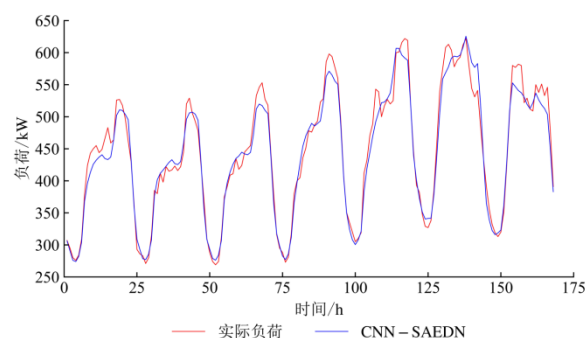


图 A7 CNN-SAEDN 的周负荷预测结果
Fig.A7 Weekly load forecasting results of CNN-SAEDN

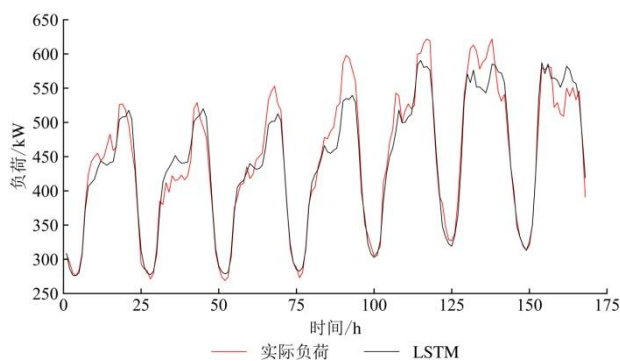


图 A8 LSTM 的周负荷预测结果
Fig.A8 Weekly load forecasting results of LSTM

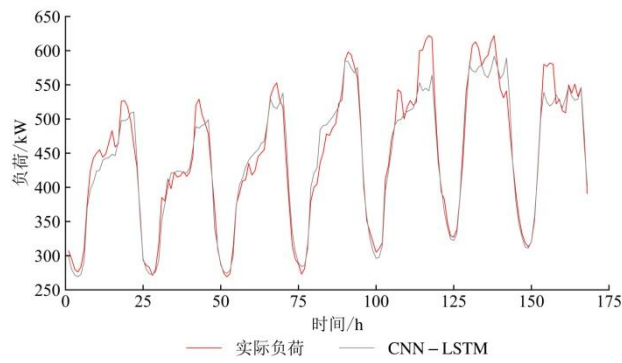


图 A9 CNN-LSTM 的周负荷预测结果
Fig.A9 Weekly load forecasting results of CNN-LSTM

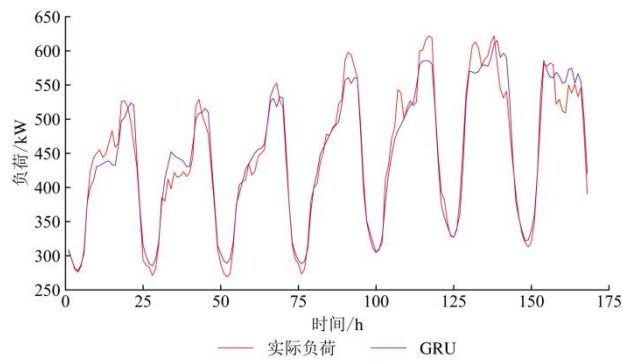


图 A10 GRU 的周负荷预测结果
Fig.A10 Weekly load forecasting results of GRU

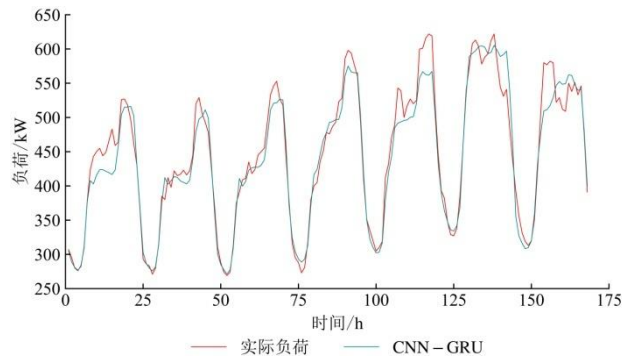


图 A11 CNN-GRU 的周负荷预测结果
Fig.A11 Weekly load forecasting results of CNN-GRU

表 A3 本文模型与注意力模型的预测结果比较

Table A3 The prediction results of the model in this paper are compared with those of the attention model

方法	第 1 天		工作日		休息日		一周	
	A_d /%	E_d /kW	A_d /%	E_d /kW	A_d /%	E_d /kW	A_d /%	E_d /kW
CNN-SAEDN	95.79	14.42	96.49	12.89	95.73	17.65	96.27	14.25
CNN-SA	95.84	13.15	95.65	14.88	94.56	20.86	95.34	16.59
CNN-LSTM-SA	95.92	14.63	95.94	14.05	95.45	19.44	95.80	15.59
CNN-GRU-SA	95.81	14.51	95.85	14.67	94.97	20.02	95.59	16.20
CNN-CA	95.79	12.75	95.41	15.84	95.00	18.40	95.30	16.57
CNN-LSTM-CA	95.17	15.97	95.76	15.21	95.34	17.64	95.64	15.90
CNN-GRU-CA	95.92	14.60	95.76	15.08	95.57	17.37	95.70	15.73

表 A4 我国南方某地区电力负荷数据集上的模型预测结果比较
Table A4 Comparison of model forecasting results on a power load dataset of a region in southern China

方法	一周		计算速度/ms
	A_d /%	E_d /MW	
CNN-SAEDN-Res	97.56	117.69	19.1
CNN	95.85	208.13	3.7
LSTM	95.76	224.37	2.9
CNN-LSTM	96.78	163.29	4.6
GRU	94.49	265.24	2.3
CNN-GRU	96.46	180.14	4.1
CNN-SA	96.37	190.94	5.9
CNN-LSTM-SA	97.15	143.46	7.1
CNN-GRU-SA	97.04	144.68	6.7
CNN-CA	96.30	181.61	5.3
CNN-LSTM-CA	97.19	141.98	6.3
CNN-GRU-CA	96.97	151.10	6.1