

基于差异补偿和短期采样对比损失的城市电力负荷预测方法

陈润桓, 戴华, 郑桂能, 李惠, 杨庚

引用本文

陈润桓, 戴华, 郑桂能, 李惠, 杨庚. 基于差异补偿和短期采样对比损失的城市电力负荷预测方法[J]. 计算机科学, 2024, 51(4): 158-164.

CHEN Runhuan, DAI Hua, ZHENG Guineng, LI Hui, YANG Geng. [Urban Electricity Load Forecasting Method Based on Discrepancy Compensation and Short-term Sampling Contrastive Loss](#) [J]. Computer Science, 2024, 51(4): 158-164.

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)

[结合卷积神经网络与多层感知机的渐进式多阶段图像去噪算法](#)

Progressive Multi-stage Image Denoising Algorithm Combining Convolutional Neural Network and Multi-layer Perceptron

计算机科学, 2024, 51(4): 243-253. <https://doi.org/10.11896/jsjcx.230100140>

[基于GAANET的立体匹配算法](#)

Algorithm of Stereo Matching Based on GAANET

计算机科学, 2024, 51(4): 229-235. <https://doi.org/10.11896/jsjcx.230100137>

[基于Transformer的街道停车位数据补全和预测](#)

Data Completion and Prediction of Street Parking Spaces Based on Transformer

计算机科学, 2024, 51(4): 165-173. <https://doi.org/10.11896/jsjcx.221200171>

[CheatKD:基于毒性神经元同化的知识蒸馏后门攻击方法](#)

CheatKD: Knowledge Distillation Backdoor Attack Method Based on Poisoned Neuronal Assimilation

计算机科学, 2024, 51(3): 351-359. <https://doi.org/10.11896/jsjcx.221200035>

[基于依赖类型剪枝的双特征自适应融合网络用于方面级情感分析](#)

Dual Feature Adaptive Fusion Network Based on Dependency Type Pruning for Aspect-based Sentiment Analysis

计算机科学, 2024, 51(3): 205-213. <https://doi.org/10.11896/jsjcx.230100035>

基于差异补偿和短期采样对比损失的城市电力负荷预测方法

陈润桓¹ 戴华^{1,2} 郑桂能³ 李惠¹ 杨庚^{1,2}

1 南京邮电大学计算机学院、软件学院、网络空间安全学院 南京 210023

2 江苏省大数据安全与智能处理重点实验室 南京 210023

3 犹他州立大学计算机科学系 洛根 84322

(1221045632@njupt.edu.cn)

摘要 城市电力负荷预测是城市智能电网规划和调度的一项重要内容。然而,城市电力负荷预测中存在数据不均的问题,给城市电力负荷预测带来了巨大挑战。传统的基于单一模型的方法难以解决数据不均的问题,而现有的基于多模型的预测方法根据电力负荷分布将数据集拆分成多个子数据集,然后分别建立多个预测模型进行预测,该类方案在一定程度上解决了数据不均问题,但存在模型构建成本较高、不同分布样本间共有的电力分布特征发生分离等问题。基于此,提出了一个轻量级城市电力负荷预测模型(Lighten-DCSC-LSTM)。该模型通过在长短期记忆网络的基础上引入差异补偿的思想和短期采样对比损失进行构建,同时构建共享特征提取层来降低模型构建成本。其中,差异补偿思想通过学习不同电力负荷分布样本之间的差异对主序列预测模块的预测结果进行差异补偿,短期采样对比损失通过动态类中心的对比学习损失对模型的训练进行正则化。为了验证模型的性能,进行了参数调优和对比实验。对比实验结果表明,模型在预测电力负荷的任务中取得了良好的性能。

关键词: 电力负荷预测;长短期记忆网络;深度学习;对比学习损失

中图分类号 TP183

Urban Electricity Load Forecasting Method Based on Discrepancy Compensation and Short-term Sampling Contrastive Loss

CHEN Runhuan¹, DAI Hua^{1,2}, ZHENG Guineng³, LI Hui¹ and YANG Geng^{1,2}

1 School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

2 Jiangsu Key Laboratory of Big Data Security & Intelligent Processing, Nanjing 210023, China

3 Computer Science Department, Utah State university, Logan 84322, USA

Abstract Urban power load forecasting is an important content of urban smart grid planning and scheduling. However, the problem of data imbalance in urban power load forecasting poses a great challenge to urban power load forecasting. Traditional single-model-based methods can hardly solve the problem of data imbalance. The existing multi-model-based forecasting methods split the datasets into multiple sub-datasets according to the electricity load profiles, and then build multiple forecasting models for forecasting, which can solve the data imbalance problem to a certain extent, but there are problems such as high model construction cost and separation of the common electricity distribution characteristics among different distribution profiles. Based on this, this paper proposes a lighten urban electric load forecasting model(Lighten-DCSC-LSTM). It is constructed by introducing the idea of discrepancy compensation and short-term sampling contrastive loss on the basis of long and short-term memory networks, while building a shared feature extraction layer to reduce the model construction cost. Among them, the idea of discrepancy compensation compensates the prediction results of the main sequence prediction module by learning the differences between different power load distribution samples, and the short-term sampling contrastive loss regularizes the training of the model by the contrastive learning loss of the dynamic class center. To verify the performance of the proposed model, parameter tuning and comparison experiments are conducted. The results of the comparison experiments show that the model achieves good performance in the task of forecasting electricity loads.

Keywords Electricity load forecasting, Long-short term memory networks, Deep learning, Contrastive learning loss

到稿日期:2023-01-17 返修日期:2023-04-21

基金项目:国家自然科学基金(61872197,61902199,61972209);安徽省高等学校科研计划重大项目(2022AH040148);江苏省研究生科研创新项目(SJCX22_0265)

This work was supported by the National Natural Science Foundation of China(61872197,61902199,61972209), Major Program of Natural Science Research Foundation of Anhui Provincial Education Department(2022AH040148) and Jiangsu Province Postgraduate Scientific Research Innovation Program(SJCX22_0265).

通信作者:戴华(daihua@njupt.edu.cn)

1 引言

随着科学技术以及城市智能电网的发展,输变电设备的智能化控制以及城市电网发电策略的快速转变已经成为可能。城市电力负荷预测是城市智能电网决策的重要工具^[1],城市智能电网通过电力负荷预测模型预测未来可能的城市用电负荷趋势,针对未来用电负荷趋势规划城市的电力调度策略和发电规模,以保持稳定的城市供需平衡关系,减少电力损失。因此,准确的城市电力负荷预测对保证城市电力系统的稳定性和经济性起着重要作用。

在城市用电负荷数据中,由于城市中的企业在周日选择停产减产,造成了周日的城市电力负荷规模普遍低于工作日的电力规模,进而导致了城市电力负荷预测中的数据不均问题。现有的基于多模型的解决方案^[2-3]是将电力负荷数据集按照不同负荷分布进行拆分,然后分别对多个子电力负荷数据集建立预测模型并进行预测。然而,这种基于多模型的方法增加了模型的构建成本,并且分离了不同负荷分布的电力负荷样本间共有的电力分布特征。

为了解决基于单一模型的方法在城市电力负荷预测任务中预测效果不佳的问题,本文提出了一种轻量级城市电力负荷预测模型(Lighten Discrepancy Compensation and Short-term Sampling Contrastive Loss LSTM, Lighten-DCSC-LSTM),该模型在长短期记忆网络的基础上引入了差异补偿思想和短期采样对比损失进行构建。其中,差异补偿思想通过构建差异补偿模块来学习不同分布电力负荷样本之间的差异,并根据预测样本的样本特征对主序列预测模块的预测结果进行差异补偿。该思想使得主序列预测模块可以更加关注学习不同电力负荷分布样本共有的电力负荷变化规律。短期采样对比损失是基于动态质心的模型训练正则化函数,它通过使一周中同一天的电力负荷序列相互接近,来确保模型注重类内的紧凑性。该损失可以使得具有相似电力负荷分布的工作日也能够被模型区分开来。同时,短期采样对比损失也增加了少样本的学习权重,从而能够更好地训练差异补偿模块。为了进一步减少模型的参数量,我们引入了共享特征提取层的概念,通过让主序列预测模块和差异补偿模块共享同一特征提取层,来减少模型的参数量。为了验证模型的性能,设计了超参数调优以及对比实验。实验结果表明,通过增加差异补偿模块和短期采样对比损失,所提模型有效地学习到了周末和工作日之间的电力负荷差异,模型在预测方面明显超越了现有的最先进的模型。

本文的主要贡献如下:

1)在长短期记忆网络中引入了差异补偿思想,以改善基于单一模型的方法在存在数据不均问题的城市电力负荷预测中不能有效预测的问题。

2)提出了短期采样对比损失作为模型训练的正则化函数。该损失使模型能够区分出具有相似负荷曲线的工作日电力负荷数据,同时增加了少样本的学习权重。

3)本文通过构建共享特征提取模块,使主序列预测模块和差异补偿模块共享同一特征提取层,在保证模型预测性能的基础上进一步减少模型的参数量。

4)Lighten-DCSC-LSTM模型与现有的各种方法进行了

实验比较。本文使用RMSE,MAPE和MAE作为实验的误差评价指标。实验结果表明,Lighten-DCSC-LSTM的性能优于现有模型。

2 相关工作

2.1 传统的电力负荷预测模型

近年来,统计方法和人工智能计算在电力预测任务中得到了广泛应用。经常用于电力负荷预测的应用统计预测模型有自回归模型(Autoregressive Model)^[4]、移动平均模型(Moving Average Model)^[5]、自回归移动平均模型(Autoregressive Moving Average Model)^[6]和差分整合移动平均自回归模型(Autoregressive Integrated Moving Average Model)^[7]。这些模型没有很高的计算要求,但当它们与基于人工智能的模型进行比较时,这些模型的预测精度较低^[8]。另一方面,人工智能技术,如人工神经网络(Artificial Neural Networks)^[9]和支持向量回归(Support Vector Regression, SVR)^[10],尽管显示出了更准确的性能,但计算成本也更昂贵。更重要的是,上述算法都没有考虑连续电力负荷系列数据之间的相关性。电力负荷数据作为一个经典的时间序列数据,连续的电力负荷时间序列之间存在着高度的相关性。因此,传统的时间序列预测方法忽略了时间维度上具有距离特征的连续电力负荷数据之间的关联性,导致预测精度有限。

2.2 基于深度学习的电力负荷预测模型

近年来,随着机器学习和深度学习的发展,研究人员针对电力负荷预测问题提出了各种深度学习模型。20世纪80年代初,Hopfield^[11]提出了递归神经网络(RNN)。RNN利用节点的顺序连接形成一个有向图,这意味着RNN有能力考虑时间序列中关于时间维度的相关性。然而,RNN在训练过程中是一个无限的动态循环过程,这可能导致梯度爆炸和梯度消失。为了解决梯度消失问题,Hochreiter等^[12]提出了一种基于RNN网络的长短期记忆网络(LSTM)。一个常规的LSTM包含一个存储单元、一个输入门、一个遗忘门和一个输出门,使得网络避免了梯度爆炸和梯度消失问题。双向长短期记忆网络(Bidirectional LSTM, Bi-LSTM)^[13]是对LSTM网络的进一步改进,它分别从前到后和从后到前学习输入的时间序列。这种学习特征对于依赖历史数据预测电力负荷的任务来说特别重要。Alhussein等^[14]提出了一个基于卷积神经网络(CNN)和长短期记忆(LSTM)相结合的深度学习框架。CNN-LSTM模型使用CNN层从输入数据中提取特征,LSTM层用于序列学习。Attention-LSTM^[15-16]将注意力机制添加到LSTM网络中,增强了LSTM网络对于长输入序列中的关键特征的学习能力。该模型也被应用于电力负荷预测,以提高预测精度^[17]。目前,这些网络在自然语言处理^[18]、语音识别^[19]、金融^[20]和风速预测^[21-23]领域很受青睐。

在关于数据不均的电力负荷预测的研究工作中,Penya等^[2]提出根据电力负荷数据的负荷情况将数据集分成3类:工作日、周六和周日。然后,他们根据一周中的不同日子构建了3种预测模型。Mansoor等^[3]通过将电力负荷数据集划分为工作日、周末和节假日3个不同的聚类,分别进行预测。在上述的研究工作中,我们发现大多数关于数据不均电力负荷预测的研究工作都采用了分割不同负荷分布样本的方法进行

分别预测。很少有人尝试在不改变原始数据集的情况下通过改进基于单一模型的方法来解决这个问题。

因此,本研究尝试在 LSTM 网络中加入学习不同负荷特

征样本差异的差异补偿模块和训练正则化的短期采样对比损失,使得改进后的模型在数据不平衡的情况下表现良好。同时,提取共享特征提取层以构建更加轻量化的电力预测模型。

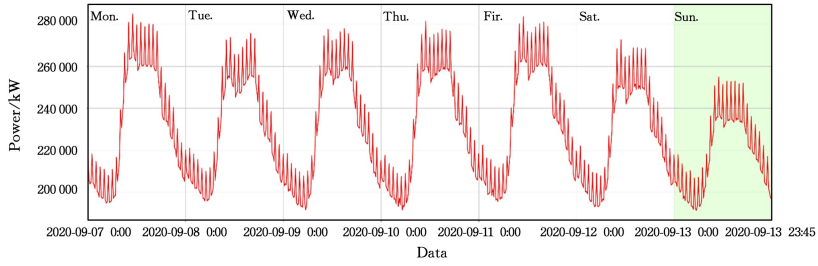


图 1 某一周的城市电力负荷分布

Fig. 1 Urban electricity load distribution in a given week

3 符号定义与问题描述

3.1 符号定义

本文使用的主要符号如表 1 所列。

表 1 符号定义

Table 1 Symbol definition

符号	定义
$X = \{X_{t-\beta}, \dots, X_{t-2}, X_{t-1}\}$	前 β 天的电力负荷数据
is_few_t	第 t 天是否是少样本
Y_t^{pre}	第 t 天的电力负荷观测结果
Y_t^{true}	第 t 天的电力负荷检测结果
H_t	第 t 日的动态类中心

3.2 问题描述

城市用电负荷主要由城市居民用电负荷和企业用电负荷组成。其中,城市居民用电负荷主要包括照明、洗衣机、空调和电视等普通家用电器的用电量,其与居民的日常生活息息相关,因此在时间维度上呈现稳定的日规律性。工业用电负荷指企业用于工业生产的用电负荷。虽然工业用电负荷根据城市居民上班时间呈现规律性,但是由于周日大部分企业选择停产减产,导致工业用电负荷在工作日与周末这两个时间段呈现较大差异,进而造成了城市电力负荷预测中的数据不均问题,给城市电力负荷的准确预测带来巨大的挑战。

本文通过对实验数据进行可视化,来具体地描述城市电力负荷数据中的日规律性和数据不均问题。

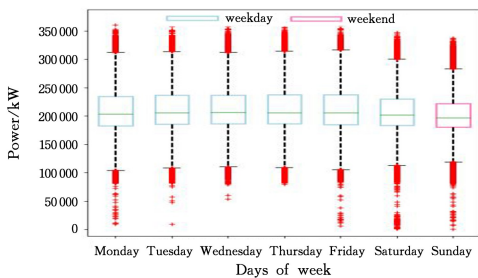


图 2 2018—2021 年的各个星期类型负荷数据分布情况箱型图

Fig. 2 Box plot of load data distribution by week type in 2018—2021

通过图 1 和图 2 可以清晰地观察到电力负荷分布中存在的两个特征:

1) 电力负荷数据存在日规律性。通过观察图 1, 可以看到电力负荷以 96 个采样点(一天)为周期呈现规律性。本文对整个实验数据集进行遍历, 电力负荷数据确实显示出很强的日规律性。

2) 城市电力负荷中周日电力负荷规模普遍低于工作日的电力负荷规模。通过观察图 2 可知, 周末的城市电力负荷规模相对于工作日有所下降, 周日的城市电力负荷规模下降更为明显。出现这种情况的原因是, 大多数企业在周日选择停止生产或降低产量, 因此周末全天的电力负荷会相对减少。同时, 因为周六的城市电力负荷规模较工作日下降幅度较小, 所以本文把周六归类为工作日。

城市电力负荷数据中的日规律性使得研究人员可以通过构建基于机器学习的电力负荷预测模型来对未来的城市电力负荷趋势进行预测。但是, 城市电力负荷中存在的非均问题, 给基于机器学习的电力负荷预测带来了挑战。

4 模型构建

4.1 长短期记忆网络

长短期记忆(LSTM)是循环神经网络(RNN)体系结构的扩展, 该体系结构在文献[12]中首次被提出, 用于解决 RNN 的梯度消失和爆炸问题。在序列建模预测任务中, 以 LSTM 为代表的循环神经网络具有较强的时间序列型数据特征提取能力。图 3 给出了 LSTM 的架构。

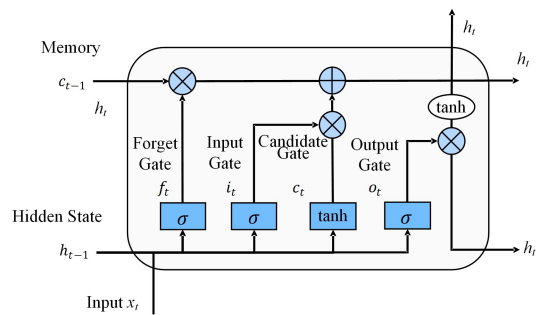


图 3 LSTM 网络结构

Fig. 3 LSTM network architecture

如图 3 所示, 每一个 LSTM 单元中均包含多个运算, 具体如式(1)~式(6)所示:

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \quad (2)$$

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (3)$$

$$\tilde{c}_t = \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (5)$$

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

其中, W_x 和 b 分别表示权重矩阵和偏置向量, f_t , i_t 和 o_t 分别为遗忘门(Forget Gate)、输入门(Input Gate)和输出门(Output Gate),这3个门的计算结果取决于当前的输入 x_t 和之前的输出 h_{t-1} 。 c_t 是单元状态(Cell State), \tilde{c}_t 是单元状态的新候选值, h_t 是时间 t 的隐藏输出(Hidden Output), \odot 是点对点的乘积。

在 LSTM 中,遗忘门决定之前被允许通过状态 h_{t-1} 的数量,以及被遗忘的数量,输入门决定新的信息哪些用来更新单元状态 c_t ,输出门决定单元状态 c_t 中哪些信息需要输出。这3个门协同工作,学习和存储时序信息。而权重 W_x 和偏差 b 是在模型训练过程中通过最小化 LSTM 单元的输出和实际训练样本之间的差异来学习到的。

4.2 短期抽样对比损失

在城市用电负荷序列中,即使是相邻工作日的用电负荷序列也会略有不同。在每天的电力负荷预测任务训练中,模型需要输出未来一天的96个观测值(15分钟的采样步骤)。这使得模型在计算训练损失时,很容易忽略同一工作日但不同日期(如周一和周二)的日用电力负荷序列中各个观测值的差异,因此有必要增加基于训练的正则化损失函数。

然而,由于电力负荷数据是一个随时间不断变化的时间序列,无法通过固定类中心使同类样本相互接近。因此本文放弃了固定类中心的方法,提出了短期采样的对比学习损失。这种损失用一个动态的类中心代替了固定的类中心。

本研究以当前预测序列的前七天电力负荷序列为类中心进行抽样。例如,当前预测的电力负荷序列的日期是星期一时,我们采样上周一的电力负荷序列作为动态类中心来计算短期抽样对比学习损失。

如图4所示,通过添加对比学习损失函数,可以鼓励来自同一类的电力负荷序列相互接近,从而确保模型更加关注类内紧凑性(Intra-class Compactness)。同时,增加该损失函数

既可以增加少样本(周末样本)的学习权重,又可以使得模型学会区别分布相似的工作日电力负荷数据。

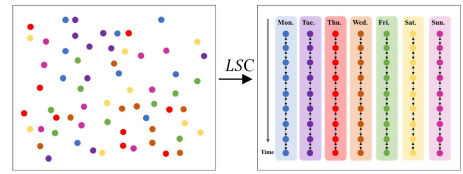


图4 短期对比损失使模型关注类内紧凑性

Fig. 4 Short-term contrast loss makes the model focus on intra-class compactness

具体的比较学习损失计算式如式(7)所示:

$$L_{SC} = \sum_{i=1}^N d_E(H_i, Y_i^{pre}) = \sum_{i=1}^N \sqrt{(H_i - Y_i^{pre})^2} \quad (7)$$

其中, N 是批处理的大小, d_E 是欧氏距离^[23]。 $H = \{H_1, H_2, \dots, H_N\}$ 和 $Y^{pre} = \{Y_1^{pre}, Y_2^{pre}, \dots, Y_N^{pre}\}$ 分别为动态类中心和批次样本的预测序列。

4.3 差异补偿与共享特征提取

本文基于差异补偿的思想将传统单一的序列预测分割成主序列预测模块(Main Sequence Forecast Module, MSF-Module)与差异补偿模块(Discrepancy Compensation Module, DC-Module),主序列预测模块和差异补偿模块均由 LSTM 与全连接(FC)层组成。如图5所示,两个模块的主要功能如下。

1) 序列预测模块:主序列预测模块忽略了不同电力负荷曲线的电力数据样本之间的差异,并对正常样本的电力负荷数据进行预测学习。

2) 差异补偿模块:差异补偿模块学习少数样本的电力负荷与正常样本的差异部分。当预测样本是少数样本时,它就会发挥作用。

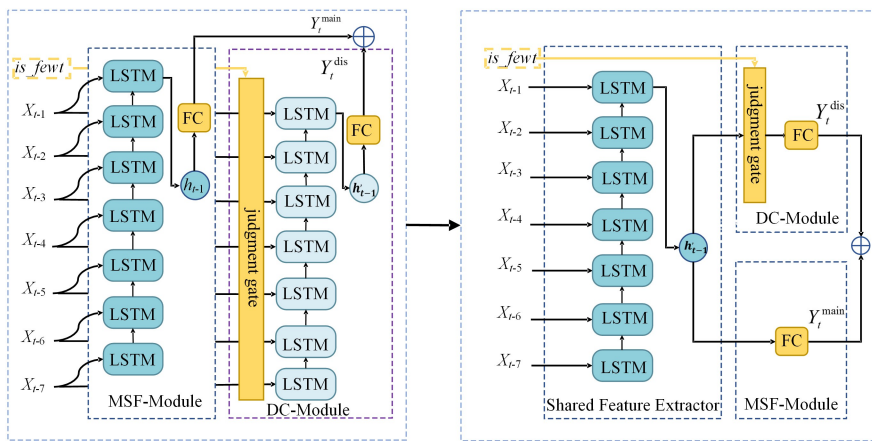


图5 提取共享特征提取层

Fig. 5 Shared feature extraction layer

在此基础上,本文通过构建共享特征提取层,将主序列预测模块和缺失填补模块中负责提取电力负荷时序特征的长短期记忆网络(LSTM)提取出来,构建一个共享特征提取器(Shared Feature Extractor),使得模型的参数量减少,同时使模型更加注重提取不同电力负荷分布样本间共有的时序特征。

4.4 Lighten-DCSC-LSTM 模型

Lighten-DCSC-LSTM 主要由共享特征提取器(Shared Feature Extractor)、主序列预测模块(MSF-Module)、差异

补偿模块(DC-Module)和训练损失计算模块(TLC-Module)组成,如图6所示。

图6中,训练损失计算模块(TLC-Module)包括MES损失和短期采样比较(Short-term Sampling Contrastive, SC)损失。MSE损失是由预测值和观测值的损失计算的,短期采样比较损失是由动态中心和观测值的损失计算的。Lighten-DCSC-LSTM 模型的具体实现步骤如下:

Step1 将输入序列 X 输入到共享特征提取层进行特征

提取,得到时序特征 F_t 。

Step2 将提取出的时序特征 F_t 输入到主序列预测模块进行预测,得到主预测序列 Y_t^{main} 。

Step3 将提取出的时序特征 F_t 和辅助输入 is_few_t 输入到 DC-Module。DC-Module 中的判断门通过 is_few_t 确定预测的样本是否是少样本。如果预是,则激活 DC-Module,对时

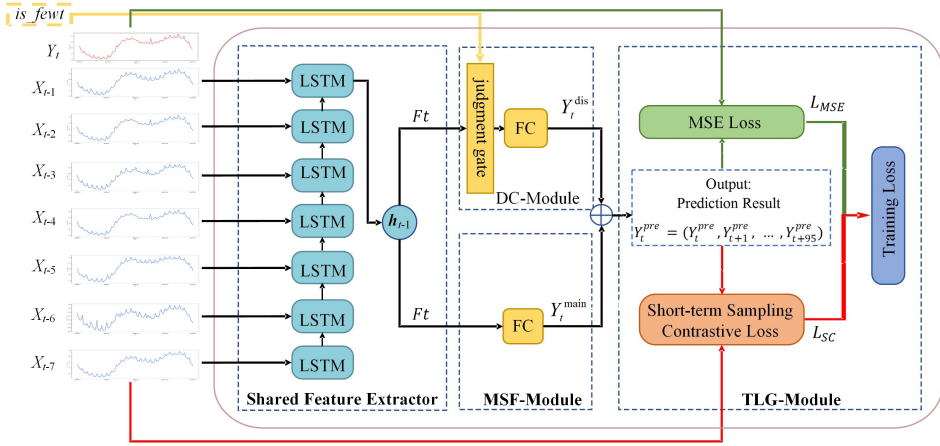


图6 Lighten-DCSC-LSTM 模型架构

Fig. 6 Lighten-DCSC-LSTM model architecture

5 实验评估

为了评估 Lighten-DCSC-LSTM 模型对数据不均的电力负荷数据的求解能力,本文进行了超参数优化和对比实验。

5.1 数据说明

本文使用的原始数据集取自第十届泰迪杯数据挖掘挑战赛 B 题。该数据集涵盖了从 2018 年 1 月 1 日 00:00 到 2020 年 8 月 31 日 23:45,为期 44 个月的时间区间。其中,城市电力负荷数据集中每隔 15 分钟采集一次城市电力负荷数据。

5.2 数据准备

由于电力负荷数据的检测和传输存在潜在的不稳定性,导致数据集中可能存在缺失、异常和噪声数据。因此在构建模型之前需要对数据集进行预处理。本研究对数据集分别进行了缺失值填补、重复值去除、离群点处理和规范化处理。我们将数据集最后 10 天和最后 30 天的城市电力负荷数据拆开作为实验的两个验证集。然后按 7:3 的比例分割其余的数据集,即训练集从 2018 年 1 月 1 日 0:00 到 2020 年 7 月 25 日 11:45,测试集从 2020 年 7 月 25 日 0:00 到 2021 年 7 月 31 日 11:45。同时,我们从需要预测的电力负荷序列 Y_t^{pre} 中提取星期类型特征,得到辅助输入 is_few_t 。

5.3 评价指标

为了衡量不同电力负荷预测算法的性能,本文使用 3 种主流的时间序列预测评价指标,包括平均绝对误差(MAE)、平均绝对百分比误差(MAPE)和均方根误差(RMSE)。

MAE 表示估计值与实际值之间总绝对误差的平均值,如式(9)所示:

$$MAE = \frac{1}{n} \sum_i |x_i - \hat{x}_i| \quad (9)$$

其中, x_i 和 \hat{x}_i 是第 i 天的实际和预测的电力负荷结果

序特征 F_t 进行差异补偿序列预测,得到 Y_t^{dis} 。然后将 Y_t^{dis} 应用于主预测序列 Y_t^{main} ,得到最终序列预测结果 Y_t^{pre} 。

Step4 将两个损失相加,得到最终的训练损失 L 。损失计算式如式(8)所示:

$$L = L_{\text{MSE}}(Y_t^{\text{true}}, Y_t^{\text{pre}}) + \lambda L_{\text{SC}}(X_{t-7}, Y_t^{\text{pre}}) \quad (8)$$

其中, λ 表示短期采样对比损失的权重。

(一天 96 个点)。

MAPE 与 MAE 类似,但为了考虑预测值与实际值之间的相对差距,MAPE 中估计值和实际值之间的每个差距都被实际值所除,如式(10)所示:

$$MAPE(\%) = \frac{100}{n} \sum_i \left| \frac{x_i - \hat{x}_i}{x_i} \right| \quad (10)$$

RMSE 通过比较估计值与实际值之间的偏差来确定精度,如式(11)所示:

$$RMSE = \sqrt{\frac{\sum_i (x_i - \hat{x}_i)^2}{n}} \quad (11)$$

其中,MAE, MAPE 和 RMSE 值越低,代表该模型越准确。

5.4 超参数调优

本文模型使用深度学习模型进行构建,模型的预测结果与其参数量依赖于模型的超参数。本文选取了最常见的两个参数进行超参数优化,分别是特征提取层中 LSTM 的隐藏层 H 以及主序列预测模块和差异补偿模块中全连接层的维度 C ,模型的参数量大小主要由这两个参数决定。隐藏层 H 与全连接层的维度 C 越大,模型的参数量越大。同时,特征提取层的隐藏层 H 越大,提取出的电力负荷特征越大,主序列预测模块以及差异补偿模块中全连接层的维度 C 也会随之增加。具体对应关系如表 2 中超参数(Hyper-parameters)与参数量(Number of Params)两列所示。

本文实验基于 Pytorch 框架进行构建。模型训练过程采用 Adam^[24]对模型参数进行优化。其中,Adam 的初始学习率 $\alpha=0.001$,短期采样对比损失的权重 $\lambda=0.5$,模型的输入序列的长度 $\beta=7$ 。同时,我们在训练过程中考虑了两个回调 API,即早期停止(EarlyStopping)和回调函数 ReduceLROnPlateau,其参数 $patience=300$, $factor=0.85$ 。表 2 列出了参数调优的结果。

表2 超参数调优结果

Table 2 Result of hyperparameter tuning

Hyper-parameters	Number of Params	Future 30 days			Future 10 days		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE
$H=6, C=6$ (shared)	3840	4404.1	3082.2	0.013	3573.6	2739.7	0.013
$H=12, C=12$ (shared)	7780	4342.3	3132.5	0.013	3538.7	2619.3	0.012
$H=24, C=24$ (shared)	16510	4246.7	3082.2	0.013	3533.6	2739.6	0.013
$H=48, C=48$ (shared)	37440	4184.3	3046.8	0.013	3303.0	2539.9	0.013
$H=96, C=96$ (shared)	93120	3927.9	2825.1	0.013	3303.0	2294.0	0.013
$H=96, C=96$ (unshared)	121150	4146.6	3081.3	0.013	3543.2	2781.2	0.013

通过表2中第五行和第六行可以看到,在其他超参数不变的情况下,通过构建共享特征提取层,可以使模型的参数量从121150下降到93120。且在对未来10天和未来30天的预测精度上,超参数 H 和 C 均取96时模型表现最优。因此,通过参数调优,本文将超参数 H 和 C 均设定为96。

5.5 对比实验

为了测试 Lighten-DCSC-LSTM 模型的有效性,我们将其与6个预测模型进行了比较,包括递归神经网络(RNN)、支持向量回归(SVR)、LSTM模型、Bi-LSTM模型、CNN-

LSTM模型和 Attention-LSTM模型。同时,我们还将该模型与基于多模型的方法构建的 LSTM模型、Bi-LSTM模型和 CNN-LSTM模型进行实验比较,结果如表3所列。可以看到,在数据不均的城市电力负荷数据中,基于多模型的方法建立的 LSTM模型、Bi-LSTM模型和 CNN-LSTM模型在未来10天和未来30天的城市电力负荷预测精度均优于基于单一模型的 LSTM模型、Bi-LSTM模型、CNN-LSTM模型和 Attention-LSTM模型。这表明基于多模型的方法确实可以缓解数据不均的问题。

表3 对比实验结果

Table 3 Result of contrast experiments

Type of Model	Model	Future 30 days			Future 10 days		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE
Single-model-based Model	SVR	9568.7	7216.8	0.030	8827.8	6590.5	0.030
	RNN	5858.3	4338.5	0.018	4501.3	3457.7	0.018
	LSTM	5369.3	3652.5	0.015	5693.2	4395.8	0.019
	Bi-LSTM	6785.9	5126.3	0.022	6470.6	5055.0	0.022
	CNN-LSTM	6962.3	5606.6	0.024	5102.5	4255.1	0.024
Multi-model-based Model	Attention-LSTM	6381.8	4651.6	0.019	5898.7	4288.5	0.019
	LSTM	4638.7	3184.9	0.015	4977.2	3804.8	0.015
	Bi-LSTM	5087.7	3722.9	0.016	5693.2	3995.8	0.016
Our Model	CNN-LSTM	5313.9	3751.7	0.016	5752.5	4255.6	0.016
	Lighten-DCSC-LSTM	3927.9	2825.1	0.013	3303.0	2294.0	0.013

同时,通过比较基于多模型的方法建立的 LSTM模型、Bi-LSTM模型和 CNN-LSTM模型的预测精度与本文提出的 Lighten-DCSC-LSTM模型的预测精度,可以明显地看出,本文提出的 Lighten-DCSC-LSTM模型在未来10天和未来30天的城市电力负荷预测精度上均优于基于多模型方法构建的3种模型。

为了更直观地展示 Lighten-DCSC-LSTM模型在存在数据不均的城市电力负荷数据集上的性能,如图7所示,我们对 Lighten-DCSC-LSTM模型在未来30天的电力负荷预测值和数据集上的真实值进行可视化(图7中,橘红色为预测值,蓝色为真实值)。可以看出,本文模型的预测结果与真实的电力负荷时间序列相似。

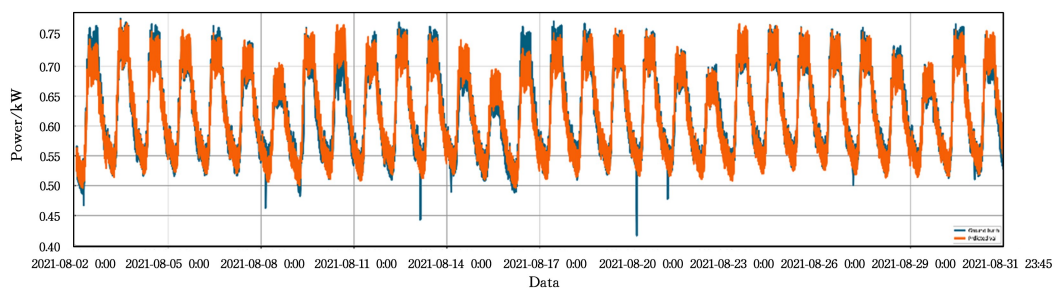


图7 模型预测结果(电子版为彩图)

Fig. 7 Model forecast results

结束语 城市电力负荷预测是城市电网中机组启停、调度和运行规划的基础。本文试图改善基于单一模型的方法在存在数据不均衡城市电力负荷数据集上表现不佳的问题。首先阐述了城市电力负荷数据集中存在数据不均问题的原因:由于企业在周日选择停产减产导致周日的电力负荷规模相比工作日明显降低,造成了城市电力负荷中的数据不均问题。

通过总结现有的电力负荷预测方法,讨论了基于多模型的方法在存在数据不均的电力负荷数据集上的不足,以及基于单一模型的方法可以更好地学习到不同样本间共有的时序特征。

因此,本文在基于单一模型方法的基础上,提出了差异补偿的思想与短期采样对比损失来解决城市电力负荷预测中的数据不均问题。同时,通过提取共享特征提取层进一步对

模型的参数量进行压缩,构建出更加轻量化的模型(Lighten-DCSC-LSTM)。通过超参数调优确定模型的超参数,然后通过对比实验证明本文模型在数据不均问题的电力负荷数据集上的表现良好,优于传统的电力预测方法以及现有的基于多模型的电力负荷预测方法。

本文工作已经证明,通过对基于单一模型的方法进行模型改进以及损失函数改进,可以在数据不均的城市电力负荷预测任务中取得更高的性能。在未来的工作中,计划通过改进短期采样对比损失函数来提高模型的预测精度。同时,将尝试应用类似的方法来解决其他数据不均的预测任务。

参 考 文 献

- [1] SHAHIDEHPOUR M, YAMIN H, LI Z. Market operations in electric power systems: forecasting, scheduling, and risk management[M]. John Wiley & Sons, 2003.
- [2] PENYA Y K, BORGES C E, FERNÁNDEZ I. Short-term load forecasting in non-residential buildings[C]// IEEE Africon'11. IEEE, 2011: 1-6.
- [3] MANSOOR M, GRIMACCIA F, LEVA S, et al. Comparison of echo state network and feed-forward neural networks in electrical load forecasting for demand response programs[J]. Mathematics and Computers in Simulation, 2021, 184: 282-293.
- [4] KOENKER R, XIAO Z. Quantile autoregression [J]. Journal of the American statistical association, 2006, 101(475): 980-990.
- [5] SINGH S, PARMAR K S, KUMAR J, et al. Development of new hybrid model of discrete wavelet decomposition and autoregressive integrated moving average (ARIMA) models in application to one month forecast the casualties cases of COVID-19 [J]. Chaos, Solitons & Fractals, 2020, 135: 109866.
- [6] JOSHUVA A, SUGUMARAN V. A machine learning approach for condition monitoring of wind turbine blade using autoregressive moving average (ARMA) features through vibration signals: a comparative study[J]. Progress in Industrial Ecology, an International Journal, 2018, 12(1/2): 14-34.
- [7] CHEN J F, WANG W M, HUANG C M. Analysis of an adaptive time-series autoregressive moving-average (ARMA) model for short-term load forecasting[J]. Electric Power Systems Research, 1995, 34(3): 187-196.
- [8] KHAN A R, MAHMOOD A, SAFDAR A, et al. Load Forecasting and Dynamic Pricing based Energy Management in Smart Grid-A Review [C] // International Multi-topic Conference. 2015.
- [9] SAVIOZZI M, MASSUCCO S, SILVESTRO F. Implementation of advanced functionalities for Distribution Management Systems: Load forecasting and modeling through Artificial Neural Networks ensembles [J]. Electric Power Systems Research, 2019, 167: 230-239.
- [10] ZHAO H, MAGOULÈS F. A review on the prediction of building energy consumption[J]. Renewable and Sustainable Energy Reviews, 2012, 16(6): 3586-3592.
- [11] HOPFIELD J. Neural networks and physical systems with emergent collective computational abilities [J]. Proceedings of the national academy of sciences, 1982, 79(8): 2554-2558.
- [12] HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [13] SCHUSTER M, PALIWAL K K. Bidirectional recurrent neural networks[J]. IEEE transactions on Signal Processing, 1997, 45(11): 2673-2681.
- [14] ALHUSSEIN M, AURANGZEB K, HAIDER S I. Hybrid CNN-LSTM model for short-term individual household load forecasting[J]. IEEE Access, 2020, 8: 180544-180557.
- [15] CHOI M, KIM H, HAN B, et al. Channel attention is all you need for video frame interpolation [C] // Proceedings of the AAAI Conference on Artificial Intelligence. 2020: 10663-10671.
- [16] LI X, LU X L. Method for Forecasting Short-Term Power Load Based on Dual-Stage Attention Mechanism and Gated Recurrent Unit Network [J]. Computer Engineering, 2022, 48 (2): 291-296, 305.
- [17] QIN J, ZHANG Y, FAN S, et al. Multi-task short-term reactive and active load forecasting method based on attention-LSTM model[J]. International Journal of Electrical Power & Energy Systems, 2022, 135: 107517.
- [18] GHOSH S, VINYALS O, STROPE B, et al. Contextual lstm (clstm) models for large scale nlp tasks[J]. arXiv: 1602.06291, 2016.
- [19] WENINGER F, ERDOGAN H, WATANABE S, et al. Speech enhancement with LSTM recurrent neural networks and its application to noise-robust ASR [C] // International Conference on Latent Variable Analysis and Signal Separation. Cham: Springer, 2015: 91-99.
- [20] CAO J, LI Z, LI J. Financial time series forecasting model based on CEEMDAN and LSTM [J]. Physica A: Statistical mechanics and its applications, 2019, 519: 127-139.
- [21] HU Y L, CHEN L. A nonlinear hybrid wind speed forecasting model using LSTM network, hysteretic ELM and Differential Evolution algorithm [J]. Energy conversion and management, 2018, 173: 123-142.
- [22] LIANG T, ZHAO Q, LV Q, et al. A novel wind speed prediction strategy based on Bi-LSTM, MOOFADA and transfer learning for centralized control centers [J]. Energy, 2021, 230: 120904.
- [23] AGRAWAL R, FALOUTSOS C, SWAMI A. Efficient similarity search in sequence databases [C] // International Conference on Foundations of Data Organization and Algorithms. Berlin: Springer, 1993: 69-84.
- [24] KINGMA D P, BA J. Adam: A method for stochastic optimization [J]. arXiv: 1412.6980, 2014.



CHEN Runhuan, born in 1998, master. His main research interests include deep learning applications and power load forecasting.



DAI Hua, born in 1982, Ph.D. professor, Ph.D supervisor, is a member of CCF (No. 40161M). His main research interests include cloud computing security and privacy protection, and power load forecasting.