

doi: 10.11835/j.issn.1000-582X.2023.201

基于 TabNet-LSTNet 的多特征短期负荷预测

吴文辉¹, 何家峰¹, 蔡高琰², 骆德汉¹

(1. 广东工业大学信息工程学院, 广州 510006; 2. 广东浩迪创新科技有限公司, 广东 佛山 528200)

摘要: 为了挖掘负荷预测中不同输入特征的重要性, 有效处理负荷数据中的线性成分和非线性成分, 提高负荷预测的精度, 提出一种基于 TabNet 和长期和短期时间序列网络 (long and short-term temporal networks, LSTNet) 的组合负荷预测模型。通过引入自监督预训练来提高 TabNet 的预测精度, 通过训练得到输入特征的全局重要性和预测结果, 然后把重要性高的特征输入到 LSTNet 训练得出预测结果, 最后通过方差-协方差组合方法得出 TabNet-LSTNet 模型的预测结果。通过仿真分析, 与传统的长短期记忆网络 (long short-term memory, LSTM)、极端梯度提升机 (extreme gradient boost, Xgboost)、轻量级梯度提升机 (light gradient boosting machine, Lightgbm) 和其他组合模型相比较, TabNet-LSTNet 模型具有更高的精度。

关键词: 负荷预测; 特征重要性; TabNet; 自监督预训练; LSTNet

中图分类号: TM715

文献标志码: A

文章编号: 1000-582X(2024)09-129-12

Multi-featured short-term load forecasting based on TabNet-LSTNet

WU Wenhui¹, HE Jiafeng¹, CAI Gaoyan², LUO Dehan¹

(1. School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, P. R. China;

2. Guangdong Haodi Innovation Technology Co., Ltd., Foshan, Guangdong 528200, P. R. China)

Abstract: To explore the importance of different input features in load forecasting, effectively handle the linear and nonlinear components in load data, and improve the accuracy of load prediction, a combined load prediction model based on TabNet and LSTNet (long and short-term temporal networks) is proposed in this paper. First, the prediction accuracy of TabNet is improved by introducing self-supervised pre-training, and then the global importance of the input features and the prediction results are obtained by training. Next, the features with high importance are input to LSTNet, which is trained to obtain the prediction results. Finally, the prediction results of the combined model are derived using the variance-covariance combination method. Simulation analysis shows that the proposed combined model has higher accuracy compared with traditional LSTM (long and short-term memory), Xgboost (extreme gradient boost), Lightgbm (light gradient boosting machine) and other combined models.

Keywords: load forecasting; feature importance; TabNet; self-supervised pre-training; LSTNet

收稿日期: 2022-07-16 网络出版日期: 2023-03-13

基金项目: 国家自然科学基金资助项目 (61571140)。

Supported by National Natural Science Foundation of China (61571140).

作者简介: 吴文辉 (1997—), 男, 硕士研究生, 主要从事负荷预测、电力大数据分析研究, (E-mail) 2112003194@mail2.gdut.edu.cn。

通信作者: 何家峰, 男, 副教授, 硕士生导师, (E-mail) 780007271@qq.com。

短期负荷预测是电力系统运行和规划的基础。准确的负荷预测可以保证电力系统安全可靠地运行、降低发电成本和提高经济效益^[1-2]。随着国家提出“碳中和、碳达峰”的目标,负荷预测对促进电力供需平衡、构建以新能源为主体的电力系统具有重要的意义^[3]。

在现代社会中,短期负荷预测受多种外部因素的相互影响,如时间因素、气候因素、用户因素和历史负荷因素等^[4]。短期负荷预测的核心问题是如何提高输入数据的质量和预测模型的水平。

根据方法类型,短期负荷预测算法可以分为数学算法、人工智能算法和组合预测算法3大类。基于数学算法的负荷预测模型有线性回归^[5]、指数平滑^[6]和累计自回归滑动平均^[7]等,这类算法只适用于简单平滑的时间序列。由于负荷数据具有非平稳特性,数学算法模型太简单,不能有效地处理这些特征信息。人工智能技术正在蓬勃发展,已有研究将人工智能算法应用到负荷预测中^[8]。基于人工智能的负荷预测算法主要有随机森林(random forest, RF)^[9]、极端学习机(extreme learning machine, ELM)^[10]、支持向量回归(support vector regression, SVR)^[11]和长短期记忆网络(long short-term memory, LSTM)^[12]等。这些算法可以有效地从时间序列中提取特征,并提供输入和输出之间的有效连接。Huang等^[11]采用SVR解决了反向传播(back propagation, BP)神经网络的预测模型收敛速度慢、对样本的依赖性高和泛化能力较弱的缺点。但基于SVR的预测模型的应用很有限,难以处理大规模的训练样本和优化超参数。Kim等^[12]提出了不同的基于LSTM的深度学习预测框架来预测住宅负荷趋势,解决了传统循环神经网络(recurrent neural network, RNN)模型中的梯度消失和梯度爆炸问题。但当输入数据有非线性特征时,模型的预测性能会下降。

为了进一步提高预测精度,许多学者提出了混合预测模型。这种方法通过将多种算法应用于数据输入和预测模型,利用每种算法和预测模型的特点和优势,满足短期负荷预测对精度的需求。

陆继翔等^[13]构建了基于CNN-LSTM的混合神经网络,通过卷积神经网络(convolutional neural networks, CNN)提取符合的多特征向量,并经由LSTM实现负荷预测,提高了预测精度。孙超等^[14]基于双层Xgboost算法考虑多特征影响进行超短期电力负荷预测,以第1层筛选出的特征集和负荷作为负荷预测层的输入变量,改善了单层模型在特征筛选和特征间相关性挖掘等方面存在的不足。该研究没有进一步挖掘不同时刻下输入特征的重要性,导致极端情况下预测精度欠佳。陈纬楠等^[15]基于LSTM和Lightgbm组合模型进行短期负荷预测,通过结合时序特征和非连续特征提高对时间序列的感知能力。但是该组合模型无法有效获取负荷序列中的长期周期性和短期非线性信息。

通过以上研究发现,Xgboost和Lightgbm等基于梯度提升决策树(gradient boosting decision tree, GBDT)的算法未能有效地挖掘出不同时刻输入特征的重要性。因为在不同时刻,每个特征对负荷变化的影响是不同的,特征的重要性会随着时间发生改变;例如,温度在一天中的不同时刻对用电量的影响是有差异的,温度的重要性随之改变。TabNet模型是Arik等^[16]提出的基于顺序注意力机制的可解释表格学习神经网络,它在保留深度神经网络(deep neural network, DNN)端到端和表征学习的基础上,还具有数模型的可解释性和稀疏特征选择等优点。传统的LSTM模型无法有效挖掘负荷数据的周期性和短期非线性规律,而LSTNet^[17]能使用不同的时间间隔对长期和短期数据特征进行信息提取,能够有效地记住多元时间序列的长期趋势和短期变化。笔者利用TabNet模型的顺序注意力机制实现特征选择,并通过引入自监督预训练来让不同的样本选择不同的特征,获取输入特征的局部重要性和全局重要性。综合考虑时间因素、气象因素和历史负荷因素对短期负荷预测的影响,通过TabNet-LSTNet模型结合两种模型的优点,既能有效处理负荷数据中的类别特征和非连续特征,又能挖掘长期负荷序列中的内在关系,提高预测的精度,促进现代电力系统安全稳定地运行。

1 基于TabNet的负荷预测模型

1.1 TabNet模型原理

TabNet模型经由多个决策步骤的子模块构成,每个决策步骤的输入数据为相同的 $B \times D$ 特征矩阵; B 是批处理量大小, D 是特征的维度。负荷特征数据通常包括类别数据和数字数据,TabNet能直接使用原始数据,将类别特征映射为数字特征。图1显示了TabNet的编码器结构,主要由特征转换器、注意力机制转换器和屏

图4为TabNet的解码器结构,每一个决策步骤都由一个特征转换器组成。编码表示为编码器中没有经过全连接层的加和向量,经过若干个步骤的加和得到重构的特征。

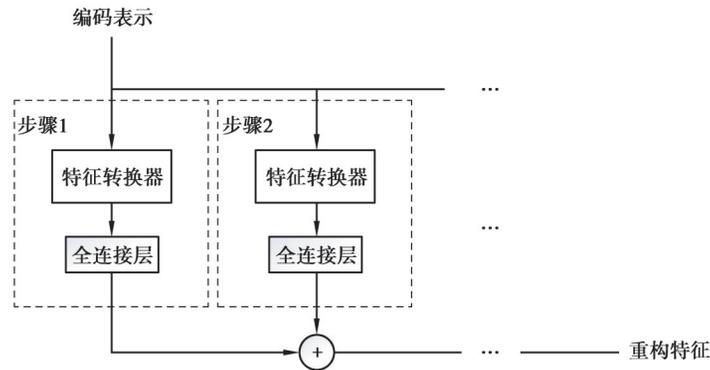


图4 TabNet解码器结构
Fig. 4 TabNet decoder architecture

1.2 TabNet负荷预测模型的自监督预训练

自监督学习主要从大量的无监督数据中挖掘监督信息,并通过构建的监督信息对网络进行训练,从而为下游任务学习提供有价值的信息。自监督预训练可以将参数约束在一个合适的空间中,使得预训练的模型能从特征空间中更合适的区域开始训练。

同一样本的不同特征是相关的,所以自监督学习时先屏蔽一些特征,如图5所示,通过上文提到的编码器-解码器结构来预测被屏蔽的特征。用这种方法训练的编码器模型可以有效地描述样本的特征。



图5 自监督预训练
Fig. 5 Self-supervised pre-training

1.3 TabNet负荷预测模型构建

选取待预测时刻的温度、湿度、降雨量、月份信息、小时信息、季度信息、工作日信息、是否月初和月末,以及前一天对应时刻的历史负荷值作为模型的输入数据,输出为待预测时刻的负荷值。然后利用训练好的模型对单个时刻分别进行预测。图6为TabNet进行负荷预测的过程。

表1为TabNet模型参数设置, N_d 为决策预测层的宽度, N_a 为注意力嵌入层的宽度, N_steps 为决策步数;这3个参数决定了模型的大小,减小这3个参数能够避免模型出现过拟合,同时不会显著降低预测精度,通过网

格搜索找出模型的最优参数。

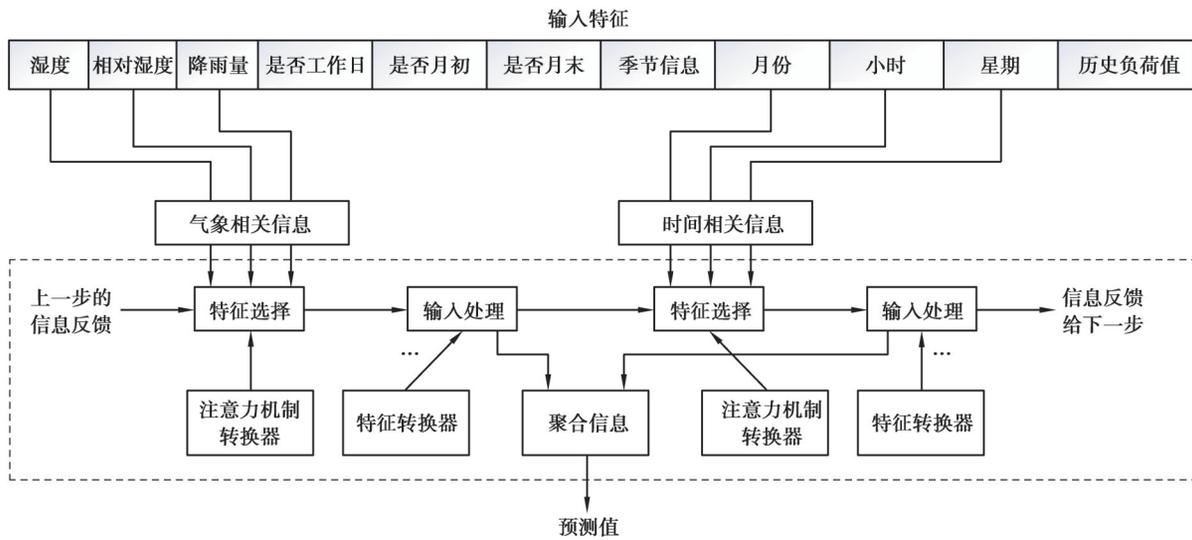


图 6 基于 TabNet 模型的负荷预测

Fig. 6 Short-term load forecasting based on TabNet

表 1 TabNet 模型参数

Table 1 TabNet model parameters

参数名	参数设置
N_d	8
N_a	8
N_steps	3
学习率	0.01
优化器	Adam

2 基于 LSTNet 的负荷预测模型

2.1 LSTNet 模型原理

LSTNet 模型^[17]的结构如图 7 所示。

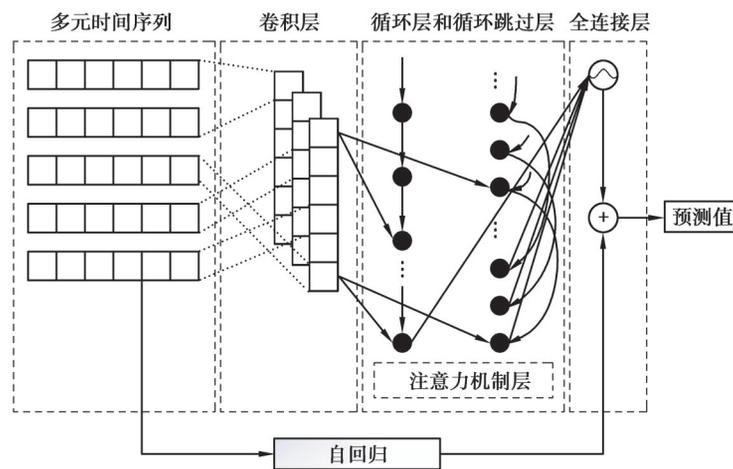


图 7 LSTNet 结构

Fig. 7 LSTNet structure

LSTNet模型使用由卷积神经网络组成的卷积层来捕捉短期特征和负荷变量的短期相关性,由长短期记忆网络组成的循环层和循环跳跃层捕捉长期特征和负荷变量的长期相关性,自适应回归部分由自回归模型组成,以提高模型的鲁棒性。

首先,数据输入到卷积层,用于提取历史负荷数据和变量之间的短期相关性。卷积层由多个滤波器组成,第 k 个滤波器由输入矩阵 \mathbf{X} 运算得到,表达式定义为

$$\mathbf{h}_k = \text{ReLU}(\mathbf{W}_k * \mathbf{X} + \mathbf{b}_k), \quad (3)$$

式中: $*$ 为卷积运算; ReLU 为激活函数; \mathbf{W}_k 为权重矩阵; \mathbf{b}_k 为偏置向量。卷积层的输出矩阵大小为 $N \times T$; N 为滤波器的数量; T 为向量的长度。

卷积层的输出作为循环层和循环跳跃层的输入变量。为了解决梯度消失问题,循环层和循环跳跃层均由LSTM组成。循环层使用激活函数更新隐藏状态。在 t 时刻,循环层的状态计算如下。

$$\mathbf{r}_t = \sigma(\mathbf{x}_t \mathbf{W}_{xr} + \mathbf{h}_{t-1} \mathbf{W}_{hr} + \mathbf{b}_r), \quad (4)$$

$$\mathbf{u}_t = \sigma(\mathbf{x}_t \mathbf{W}_{xu} + \mathbf{h}_{t-1} \mathbf{W}_{hu} + \mathbf{b}_u), \quad (5)$$

$$\mathbf{c}_t = \text{ReLU}(\mathbf{x}_t \mathbf{W}_{xc} + \mathbf{r}_t \otimes (\mathbf{h}_{t-1} \mathbf{W}_{hc}) + \mathbf{b}_c), \quad (6)$$

$$\mathbf{h}_t = (1 - \mathbf{u}_t) \otimes \mathbf{h}_{t-1} + \mathbf{u}_t \otimes \mathbf{c}_t. \quad (7)$$

式中: \otimes 表示元素之间相乘; σ 表示Sigmoid函数; \mathbf{r}_t 为LSTM的遗忘门; \mathbf{u}_t 为更新门; \mathbf{h}_t 为隐藏层输出。

LSTM无法捕获非常长期的相关性,添加了循环跳跃层来解决该问题。这是一个具有时间跳跃功能的循环结构,它扩展了输入序列的时间范围,简化了优化步骤。具体来说,在相邻时段中,在当前隐藏单元和相邻时间段的隐藏单元之间添加跳过连接。更新过程可以表示为

$$\mathbf{r}_t = \sigma(\mathbf{x}_t \mathbf{W}_{xr} + \mathbf{h}_{t-p} \mathbf{W}_{hr} + \mathbf{b}_r), \quad (8)$$

$$\mathbf{u}_t = \sigma(\mathbf{x}_t \mathbf{W}_{xu} + \mathbf{h}_{t-p} \mathbf{W}_{hu} + \mathbf{b}_u), \quad (9)$$

$$\mathbf{c}_t = \text{ReLU}(\mathbf{x}_t \mathbf{W}_{xc} + \mathbf{r}_t \otimes (\mathbf{h}_{t-p} \mathbf{W}_{hc}) + \mathbf{b}_c), \quad (10)$$

$$\mathbf{h}_t = (1 - \mathbf{u}_t) \otimes \mathbf{h}_{t-p} + \mathbf{u}_t \otimes \mathbf{c}_t. \quad (11)$$

式中, p 为跳过的隐藏单元数。

跳过的隐藏单元数需要作为超参数预先设置,这不利于提取时间序列中的非季节性特征,因为随着时间推移,周期长度也会相应地改变。为了解决这一问题,采用注意力机制来关注时间序列中重要性高的特征,为

$$\alpha_t = \text{AttnScore}(\mathbf{H}_t^R, \mathbf{h}_{t-1}^R), \quad (12)$$

式中: α_t 为 t 时刻下的注意力权重; AttnScore 为一些相似的函数,如点积、余弦; $\mathbf{H}_t^R = [\mathbf{h}_{t-q}^R, \dots, \mathbf{h}_{t-1}^R]$ 为LSTM中的隐藏状态的堆叠矩阵。时间注意力层的最终输出为

$$\mathbf{h}_t^D = W[\mathbf{c}_t; \mathbf{h}_{t-1}^R] + \mathbf{b}, \quad (13)$$

式中, \mathbf{h}_t^D 为加权向量 \mathbf{c}_t 和隐藏状态 \mathbf{h}_{t-1}^R 的串联。

通过全连接层结合循环层和循环跳跃层的输出结果。卷积层、循环层和循环跳跃层组合成一个具有非线性特性的神经网络。然而,在实际的负荷数据中,负荷数据通常在峰值附近不断变化,因此会影响预测结果。为了解决峰值不敏感的问题,在LSTNet中,使用自回归模型作为线性分量,形成自适应回归模块。自回归模型表示如下:

$$\mathbf{h}_{t,i}^L = \sum_{k=0}^{q^a-1} w_k^{\text{ar}} y_{t-k,i} + \mathbf{b}^{\text{ar}}. \quad (14)$$

式中: $\mathbf{h}_{t,i}^L$ 为AR模型的输出; w_k^{ar} 和 \mathbf{b}^{ar} 为模型参数; q^a 为输入矩阵的大小。

最后,将神经网络和自回归的结果结合起来,输出最终的预测值 \hat{Y}_t 。

$$\hat{Y}_t = \sigma(\mathbf{h}_t^D + \mathbf{h}_t^L). \quad (15)$$

2.2 LSTNet负荷预测模型构建

表2为LSTNet模型参数。历史负荷能够明显反映负荷的变化规律,因此,需要把历史负荷值作为短期负荷预测模型的输入特征。历史负荷值选取过多或过少都会影响模型的预测精度,文中时间窗口设置为4,即

选取过去 4 个时刻的负荷值作为模型的输入;由于负荷数据还具有明显的日周期性,把跳过单元数设置为 24,即选取前一天同一时刻的负荷值也作为模型的输入变量,同时获取负荷数据的短期关系和长期关系。AR 窗口尺寸设置为 7,即选取过去 7 个时刻的负荷值进行线性回归,以此来解决负荷数据在峰值处不敏感的问题。

表 2 LSTNet 模型参数
Table 2 Parameters of the LSTNet model

参数名	参数设置
CNN 滤波器数量	100
CNN 卷积核维度	9
LSTM 隐藏神经元数	64
时间窗口宽度	4
跳过单元数	24
AR 窗口尺寸	7
优化器	Adam
学习率	0.01

3 TabNet-LSTNet 负荷预测模型

3.1 输入特征选择

负荷值由许多因素共同决定,因此,有效地提取和选择输入特征是提高模型性能的重要手段。表 3 为所选用的输入特征,通过 TabNet 的顺序注意力机制来实现特征选择。

表 3 输入特征
Table 3 Input features

序号	名称	序号	名称	序号	名称	序号	名称
1	最高温度	5	降雨量	9	是否月初	13	历史负荷值
2	最低温度	6	当日小时信息	10	是否月末		
3	平均温度	7	当日星期信息	11	季节信息(春夏秋冬)		
4	相对湿度	8	当日月份信息	12	是否工作日		

图 8 显示了每个特征的局部重要性,TabNet 在第 1 个和第 3 个决策步骤之间进行特征选择,颜色越亮表示特征在这个决策步骤中被赋予更大的权重。

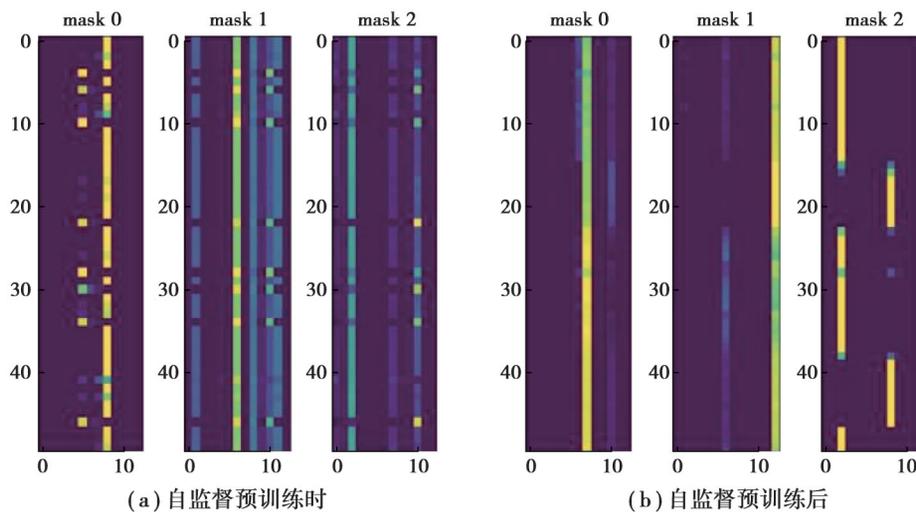


图 8 特征的局部重要性

Fig. 8 The local importance of features

图9显示了特征的全局重要性,最高温度、平均温度、历史负荷值占比较大,而最低温度、相对湿度、降雨量占比最小,所有特征的重要性相加等于1。因此,把重要性较低的2、4、5号特征删除,其余特征用作LSTNet模型的输入数据。

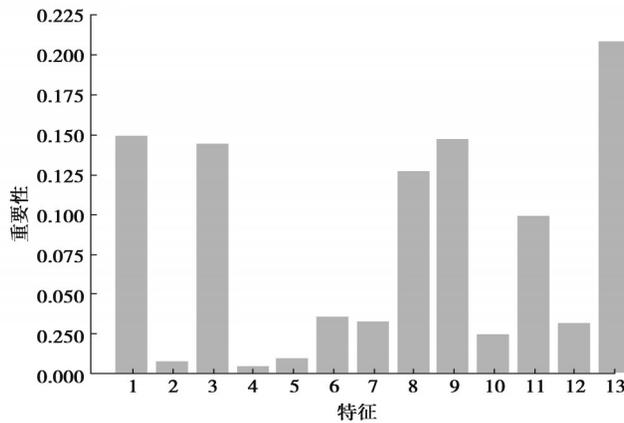


图9 特征的全局重要性

Fig. 9 Global importance of features

3.2 组合预测模型

在经过 TabNet 模型筛选特征后,通过滑动窗口构建 LSTNet 模型的输入数据,采用单步预测的方式预测当前负荷值。

通过 TabNet 和 LSTNet 分别得到预测结果后,采用方差-协方差组合预测法^[18-19]确定 2 个模型的权重系数。假设 2 个模型的预测值分别为 y_1 和 y_2 , 预测值和真实值误差分别为 e_1 和 e_2 , 方差分别为 σ_1 和 σ_2 , 组合预测结果为

$$y_c = w_1 y_1 + w_2 y_2, \tag{16}$$

$$e_c = w_1 e_1 + w_2 e_2, \tag{17}$$

$$\text{Var}(e_c) = w_1^2 \sigma_1 + w_2^2 \sigma_2, \tag{18}$$

$$w_1 + w_2 = 1. \tag{19}$$

式中: y_c 为组合预测结果; e_c 为组合预测误差; $\text{Var}(e_c)$ 为组合预测方差; w_1 和 w_2 分别为 2 个模型的权重系数。

为了求出合适的权重系数, $\text{Var}(e_c)$ 取值应最小, 可通过拉格朗日乘子法求出 w_1 和 w_2 的最小值, 如式 (20) 所示。最后, 得出组合预测结果 y_c 。图 10 为组合预测过程。

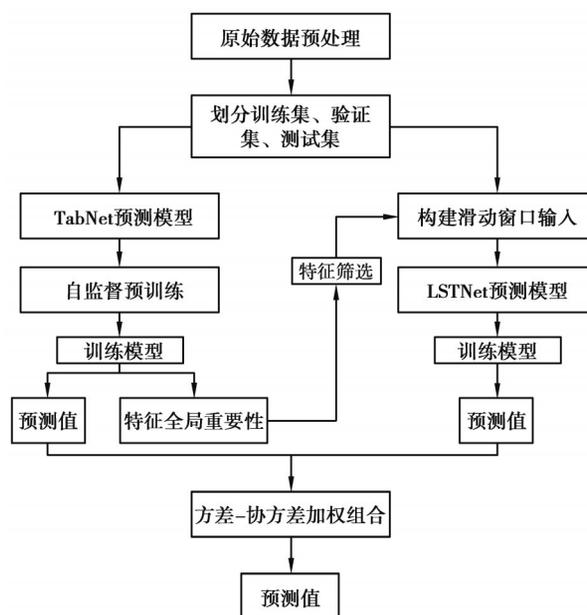


图10 组合预测流程图

Fig. 10 Flowchart of the combined prediction

$$w_1 = \frac{\sigma_2}{\sigma_1 \sigma_2}, w_2 = \frac{\sigma_1}{\sigma_1 \sigma_2} \circ \quad (20)$$

4 实验结果与分析

4.1 实验数据及预处理

选取中国广东省佛山市某地区2015年1月1日至2016年12月31日每小时的电力负荷数据、对应的气象数据和日期相关数据,每天采集24个点,一共17544条数据;按照8:1:1的比例划分为训练集、验证集、测试集。为了验证组合模型的有效性,从测试集中再划分,前50%的数据用于计算真实值和预测值的误差,为后50%的数据提供组合依据,再用这些数据验证模型的预测精度。

为了避免缺失数据对负荷预测的不利影响,并考虑负荷数据的随机性和特殊性^[20],实验复制了前一天同时观测到的数据,以填补缺失值。为了提高预测的精度和收敛速度,对输入数据和输出数据进行归一化处理:

$$\tilde{x}(i) = (x(i) - x_{\min}) / (x_{\max} - x_{\min}) \circ \quad (21)$$

式中: $\tilde{x}(i)$ 为归一化后的值; x_{\max} 和 x_{\min} 分别为原始样本数据的最大值和最小值。

4.2 评价指标

采用平均绝对百分比误差 R_{MAPE} (mean absolute percentage error, MAPE)、平均绝对误差 R_{MAE} (mean absolute error, MAE)和均方根误差 R_{RMSE} (root mean square error, RMSE)3个指标来评价模型的预测能力。

R_{MAPE} 是实际负荷值和预测负荷值的绝对误差之和除以实际值的平均值,是评估预测模型可靠性的最常用指标。 R_{MAE} 和 R_{RMSE} 用于测量预测负荷值的总体偏差,它与电力负荷具有相同的单位,使评估模型的预测能力变得方便和直接。 R_{MAPE} 、 R_{MAE} 和 R_{RMSE} 的定义如下:

$$R_{\text{MAPE}} = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - y_c|}{y_i}, \quad (22)$$

$$R_{\text{MAE}} = \frac{1}{n} \sum_{i=1}^n |y_i - y_c|, \quad (23)$$

$$R_{\text{RMSE}} = \sqrt{\frac{1}{n} \sum_{i=1}^n |y_i - y_c|^2} \circ \quad (24)$$

式中: n 为预测样本数; y_i 和 y_c 分别为 i 时刻的真实负荷值和预测负荷值。

4.3 消融实验

为了验证所提的自监督预训练和特征筛选的优越性进行了消融实验。通过表4的消融实验误差可知,TabNet经过自监督预训练后, R_{MAPE} 、 R_{RMSE} 和 R_{MAE} 分别降低了0.47%、41.20 MW、36.02 MW,证明自监督预训练能够让模型更有效地学习样本的特征,深入挖掘了不同时刻下输入特征的重要性,提高模型的性能;LSTNet经过特征筛选后, R_{MAPE} 、 R_{MAE} 和 R_{RMSE} 分别降低了0.37%、43.90 MW、35.59 MW,说明特征筛选让模型只关注重要性高的特征,避免了不相关特征的干扰,有效地提升了预测性能。

表4 消融实验误差对比

Table 4 Comparison of ablation experiment errors

模型	$R_{\text{MAPE}}/\%$	$R_{\text{RMSE}}/\text{MW}$	R_{MAE}/MW
TabNet有自监督预训练	1.37	124.96	88.94
TabNet无自监督预训练	1.84	166.16	124.95
LSTNet有特征筛选	1.75	150.38	115.86
LSTNet无特征筛选	2.12	193.57	151.45

4.4 预测结果对比分析

为了验证所提出方法的可靠性和有效性,将实验结果与其他模型的预测结果进行比较。由表5可知,

LSTNet模型比传统的LSTM网络预测效果有所提升, R_{MAPE} 、 R_{RMSE} 和 R_{MAE} 分别下降了1.64%、162.88 MW和113.03 MW;相对于Xgboost、Lightgbm、Catboost这些GBDT模型,TabNet的预测能力也有显著提升, R_{MAPE} 为1.37%,在所有单一模型中效果最好,说明TabNet模型能更好地表征负荷数据。

表5 不同模型结果对比

Table 5 Comparison of different model results

方法	$R_{MAPE}/\%$	R_{RMSE}/MW	R_{MAE}/MW
LSTM	3.39	313.26	228.89
LSTNet	1.75	150.38	115.86
Xgboost	2.64	247.05	188.83
Lightgbm	2.85	273.91	201.78
Catboost	2.57	220.94	177.15
TabNet	1.37	124.96	88.94
CNN-LSTM	2.17	195.08	151.26
Lightgbm-LSTM	2.08	183.52	143.06
Catboost-LSTNet	1.54	135.45	104.49
TabNet-LSTNet	1.19	105.29	78.94

由图11(a)可知,在单一模型中,TabNet模型和LSTNet模型能更好地拟合负荷曲线。在组合模型中,采用GBDT模型中效果较好的Catboost和LSTNet组合,分别与文献[13]中的CNN-LSTM模型、文献[15]中的Lightgbm-LSTM和文中所提出的TabNet-LSTNet模型相比较,TabNet-LSTNet模型的3项指标均为最小,分别为1.19%、105.29 MW和78.94 MW;由图11(b)可知,在负荷波动较大的峰谷和峰底附近,TabNet-LSTNet模型比其他组合模型更好地拟合负荷曲线,说明所提出模型能够充分挖掘负荷数据规律和时序信息。

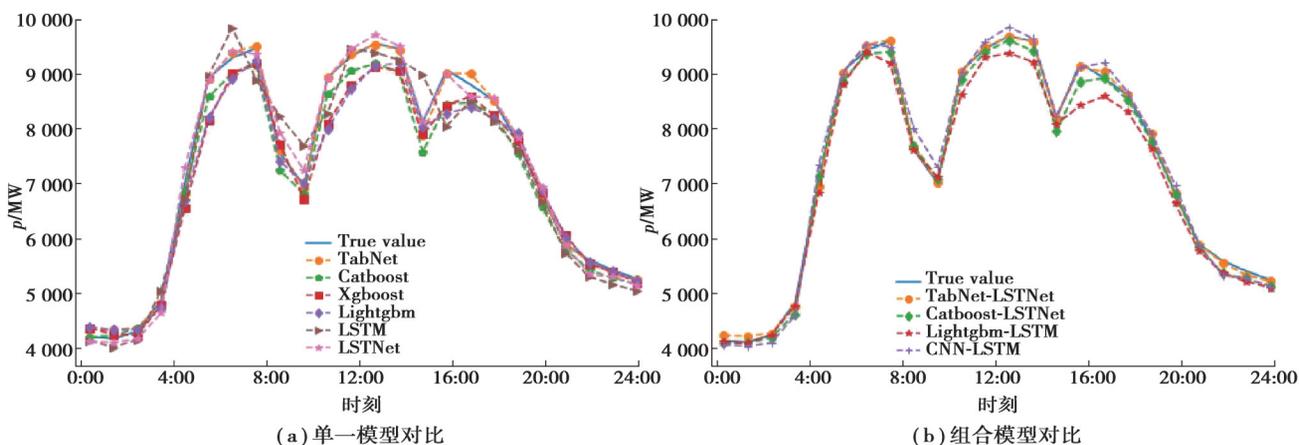


图11 不同模型对比

Fig. 11 Comparison of different models

图12为文中组合模型与单一模型各时刻的预测误差对比,与单一的TabNet模型和LSTNet模型相比,TabNet-LSTNet模型的预测误差波动更小。在随机性和波动性较强的情况下,组合模型有效地降低整体预测误差,提高了预测性能。

表6为TabNet-LSTNet模型和单一模型一天的预测结果对比,组合模型的预测误差最小, R_{MAPE} 、 R_{RMSE} 和 R_{MAE} 分别为0.62%、56.33 MW和42.31 MW。该模型充分学习了负荷序列中的依赖关系,能让供电部门更好地安排日调度计划,提高经济效益和安全保障。

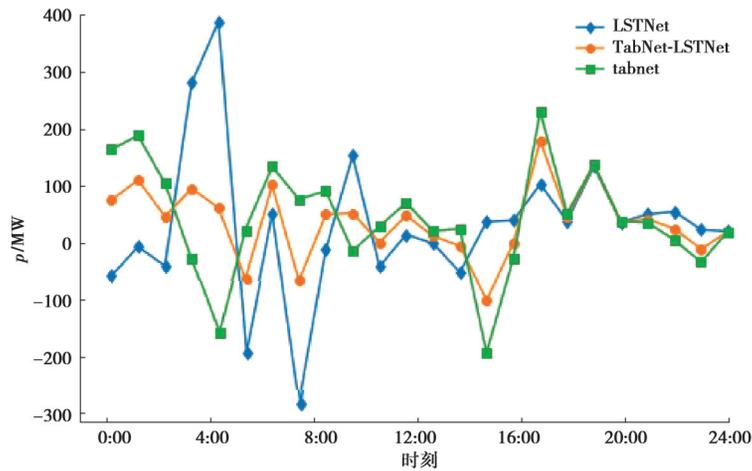


图 12 预测误差对比

Fig. 12 Comparison of forecast errors

表 6 一天的预测结果对比

Table 6 Comparison of results in one day

方法	$R_{MAPE}/\%$	R_{RMSE}/MW	R_{MAE}/MW
LSTNet	1.15	116.49	74.96
TabNet	1.01	85.33	62.07
TabNet-LSTNet	0.62	56.33	42.31

5 结束语

为了满足供电部门对负荷预测高精度的要求,提出了一种基于 TabNet-LSTNet 的短期负荷预测模型。通过实验分析,得出以下结论。

- 1) TabNet 对影响负荷预测的特征具有良好的解释性,能够挖掘出每个特征的局部重要性和全局重要性。
- 2) LSTNet 有效地处理时间序列中的线性特征和非线性特征,与传统的循环神经网络相比,能够获取相当长时间的相关信息和短期非线性变化规律。
- 3) TabNet-LSTNet 模型结合 2 种方法的优点,有效地降低了误差指标,提高了模型的整体预测性能,为电力系统的稳定运行提供更有力的保障。

在未来的研究中,可将负荷预测应用到多个具体的行业,例如商业用电、民用用电、工业用电等,挖掘出每个行业数据特征的规律,提高全行业的负荷预测准确率。

参考文献

- [1] 焦筱俊,徐青山. 电力系统单用户超短期负荷预测算法研究[J]. 电测与仪表, 2020, 57(1): 30-35, 48.
Jiao X Q, Xu Q S. A new ultra-short-term load forecasting algorithm for single user in power system[J]. Electrical Measurement & Instrumentation, 2020, 57(1): 30-35, 48.(in Chinese)
- [2] Wang Y Y, Sun S F, Chen X, et al. Short-term load forecasting of industrial customers based on SVM and XGBoost[J]. International Journal of Electrical Power & Energy Systems, 2021, 129: 106830.
- [3] López J C, Rider M J, Wu Q W. Parsimonious short-term load forecasting for optimal operation planning of electrical distribution systems[J]. IEEE Transactions on Power Systems, 2019, 34(2): 1427-1437.
- [4] Massaoudi M, Refaat S S, Chihi I, et al. A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for Short-Term Load Forecasting[J]. Energy, 2021, 214: 118874.
- [5] Fang J L, Xing Y, Fu Y, et al. Rural power system load forecast based on principal component analysis[J]. Journal of Northeast Agricultural University (English Edition), 2015, 22(2): 67-72.

- [6] Ji P R, Xiong D, Wang P, et al. A study on exponential smoothing model for load forecasting[C]//2012 Asia-Pacific Power and Energy Engineering Conference, Shanghai, China. IEEE, 2012: 1-4.
- [7] Zhu X H, Shen M. Based on the ARIMA model with grey theory for short term load forecasting model[C]//2012 International Conference on Systems and Informatics (ICSAI2012), Yantai, China. IEEE, 2012: 564-567.
- [8] 徐晴,周超,赵双双,等. 基于机器学习的短期电力负荷预测方法研究[J]. 电测与仪表, 2019, 56(23): 70-75.
Xu Q, Zhou C, Zhao S S, et al. Research on short-term power load forecasting method based on machine learning[J]. Electrical Measurement & Instrumentation, 2019, 56(23): 70-75.(in Chinese)
- [9] 蒋敏,顾东健,孔军,等. 基于在线序列极限支持向量回归的短期负荷预测模型[J]. 电网技术, 2018, 42(7): 2240-2247.
Jiang M, Gu D J, Kong J, et al. Short-term load forecasting model based on online sequential extreme support vector regression [J]. Power System Technology, 2018, 42(7): 2240-2247.(in Chinese)
- [10] Ertugrul Ö F. Forecasting electricity load by a novel recurrent extreme learning machines approach[J]. International Journal of Electrical Power & Energy Systems, 2016, 78: 429-435.
- [11] Huang N T, Lu G B, Xu D G. A permutation importance-based feature selection method for short-term electricity load forecasting using random forest[J]. Energies, 2016, 9(10): 767.
- [12] Kim Y, Son H G, Kim S. Short term electricity load forecasting for institutional buildings[J]. Energy Reports, 2019, 5: 1270-1280.
- [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 J X, Zhang Q P, Yang Z H, 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.(in Chinese)
- [14] 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 C, Lü Q, Zhu S T, 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.(in Chinese)
- [15] 陈纬楠,胡志坚,岳菁鹏,等. 基于长短期记忆网络和LightGBM组合模型的短期负荷预测[J]. 电力系统自动化, 2021, 45(4): 91-97.
Chen W N, Hu Z J, Yue J P, et al. Short-term load prediction based on combined model of long short-term memory network and light gradient boosting machine[J]. Automation of Electric Power Systems, 2021, 45(4): 91-97.(in Chinese)
- [16] Arik S Ö, Pfister T. TabNet: attentive interpretable tabular learning[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2021, 35(8): 6679-6687.
- [17] Lai G K, Chang W C, Yang Y M, et al. Modeling long- and short-term temporal patterns with deep neural networks[EB/OL]. 2017: arXiv: 1703.07015. <http://arxiv.org/abs/1703.07015>.
- [18] 朱国森. 基于Stacking与Prophet组合模型的短期电力负荷预测[D]. 淮南: 安徽理工大学, 2021.
Zhu G S. Short-term power load forecasting based on the combined model of stacking and prophet[D]. Huainan: Anhui University of Science & Technology, 2021. (in Chinese)
- [19] 唐祥玲,王平,李思岑,等. 基于方差-协方差组合预测的中长期电力负荷预测研究[J]. 电气技术, 2015(1): 15-18.
Tang X L, Wang P, Li S C, et al. Research on medium and long-term electric load forecasting based on variance-covariance combined model[J]. Electrical Engineering, 2015(1): 15-18.(in Chinese)
- [20] Wang Y Y, Chen J, Chen X Q, et al. Short-term load forecasting for industrial customers based on TCN-LightGBM[J]. IEEE Transactions on Power Systems, 2021, 36(3): 1984-1997.

(编辑 吕建斌)