

DOI:10.19651/j.cnki.emt.2211349

基于特征优选策略和 DLSTMs-FCN 优化的短期负荷预测*

孟鑫鑫 黄山 印月
(四川大学电气工程学院 成都 610000)

摘要: 针对当前基于长短期记忆网络的短期负荷预测模型存在特征冗余、重要信息丢失等问题,提出一种基于特征优选策略和 DLSTMs-FCN 并联优化结构的短期负荷预测方法。首先利用基于极限梯度提升的特征优选策略构造负荷预测模型中的输入特征最优集,减少冗余信息,加快模型拟合;而后利用 DLSTMs 提取负荷数据的时序特征,并辅以 FCN 的多维卷积运算及结构特征提取的高分辨率信息,增强对输入数据重要特征的学习和记忆,进而并联构成高效准确的短期负荷预测模型。实验结果表明,本文优化方法相较于 ALSTMs 和 CNN-LSTMs 预测误差分别降低了 6% 和 4%,预测误差波动分别降低了 4.7% 和 4.8%。

关键词: 短期负荷预测;特征优选策略;极限梯度提升;长短期记忆;FCN

中图分类号: TM715 文献标识码: A 国家标准学科分类代码: 470.40

Short-term load forecasting model based on feature optimization strategy and DLSTMs-FCN

Meng Jinxin Huang Shan Yin Yue
(School of Electrical Engineering, Sichuan University, Chengdu 610000, China)

Abstract: The short-term load forecasting model using long short-term memory(LSTM) network has the problem of feature redundancy and loss of important information. In order to solve these problems, a short-term load forecasting method based on feature selection strategy and DLSTMs-FCN is proposed. Firstly, the feature optimization strategy based on extreme gradient boosting(Xgboost) is adopted to improve the feature redundancy problem of load prediction input. Secondly, DLSTMs are used to extract the time series features of load data, and the high-resolution information is extracted through the multi-dimensional convolution operation of FCN and structural features. The purpose is to enhance the learning and memory of important features of input data, and then form an efficient and accurate short-term load forecasting model in parallel. The experimental results show that compared with ALSTMs and CNN-LSTMs, the prediction error of the optimization method in this paper decreases by 6% and 4% respectively, and the prediction error fluctuation decreases by 4.7% and 4.8% respectively.

Keywords: short-term load forecasting;feature selection strategy;Xgboost;LSTM;FCN

0 引言

短期负荷预测是电力系统运营商开展资源配置和制定调度方案的重要依据^[1]。高精度的短期负荷预测有助于电力系统降低备用容量,保障用电安全的同时提升经济效益。

以向量机^[2]、指数平滑^[3]和自回归积分滑动平均^[4]等为代表的经典负荷预测模型主要针对线性关系进行统计分析,对环境、日期等因素与负荷的非线性关系考虑不足,预

测效果较差^[5]。伴随深度学习的兴起,以长短期记忆神经网络(long short-term memory, LSTM)为代表的循环神经网络(recurrent neural network, RNN)凭借优越的时序数据建模能力在负荷预测研究中受到重视^[6-7]。文献[8]提出一种改进型黑猩猩算法优化的 LSTM 短期负荷预测模型。研究表明,随着时序建模长度增加, LSTM 会遗忘历史关键特征信息^[9]。因此,文献[10]针对 LSTM 的重要信息丢失问题,在 LSTM 模型中引入注意力机制(Attention),突出

收稿日期:2022-09-10

* 基金项目:四川省科技计划资助项目(2021YFSY0051)、四川省科技厅国际/港澳台科技创新合作项目(2022YFH0018)资助

对负荷预测具有关键影响的输入特征,有效提高了模型预测精度。文献[11-12]提出结合卷积神经网络(convolution neural network, CNN)和 LSTM 的串行负荷预测模型, CNN 提取特征为 LSTM 提供输入,使负荷预测精度和预测效率得到提升。虽然上述文献通过 Attention 或 CNN 等机制突出重要输入特征,但是所提模型均为串行结构,前向过程中由于遗忘门的存在,仍会丢失重要信息。

此外,现有信息数据存在海量高维的特点^[13],全部信息数据作为输入特征不仅增加了训练的计算开销,而且冗余特征对以 LSTM 模型为主体的负荷预测模型的预测精度会带来负面作用^[14]。现有研究通常采用相关系数^[15]或互信息^[16]作为度量指标。其中相关系数计算量小但无法挖掘非线性关系^[16];互信息能够挖掘变量间的复杂关系,但不同关联方式的健壮性不一致,存在公平性问题^[16]。除此之外,相关系数或互信息对应的选择往往通过主观给定筛选阈值,过低无法筛除冗余特征,过高易丢失关键特征。

因此,本文提出一种基于多层丢失长短时记忆与全卷积并联优化网络(dropout long short-term memorys with fully convolution neural network, DLSTMs-FCN)和特征优选策略的短期负荷预测模型。首先针对特征冗余问题,提出一种基于极限梯度提升(extreme gradient boosting, Xgboost)的特征优选策略,采用 Xgboost 计算平均特征增益作为特征重要性,充分挖掘多特征间的共同作用,并采用后向搜索策略为负荷预测模型输入端提供特征优选集合。其次针对重要信息丢失问题,提出 DLSTMs-FCN 并联网结构。在分支模块多层丢失长短时记忆网络(dropout long short-term memorys, DLSTMs)中通过 LSTMs 充分捕获特征数据流中的时序信息,为避免 LSTMs 中的过拟合问题,引入 Dropout^[17]防止模型陷入局部最优解;在分支模块全卷积网络(fully convolution network, FCN)中,通过小卷积核细化特征提取的粒度,充分捕获负荷与特征间高分辨率细节^[18]。最后在输出端通过并联结构加权融合 DLSTMs 提取的时序信息与 FCN 提取的高分辨率信息,进而取得更加准确的负荷预测结果。在公开数据集上,对本文提出的方法进行实验验证,并与常见预测模型进行性能比较。实验表明,本文提出的方法提高了短期负荷预测的精度和稳定性,相较于常见的预测模型可以更加准确的预测负荷变化。

1 基于 Xgboost 的特征优选策略

特征选择能够剔除冗余特征,从而精简特征数量,提高预测精度^[9]。在电力负荷预测研究中,通常采用相关性分析进行特征选择。然而,相关性分析忽略了特征间的耦合关系,得到的特征组合可能存在特征冗余的问题,影响模型预测结果。为减少输入特征之间的冗余,优化 LSTM 模型的输入特征,本文提出一种基于 Xgboost 的特征优选策略。通过 Xgboost 构建输入特征与目标负荷之间的回归关系,

然后将 Xgboost 的特征平均增益作为对应特征的重要程度量指标,最后通过后向搜索策略排除冗余特征。

1.1 Xgboost 算法

Xgboost 是一种基于 Boosting 原理的集成学习算法。相较于自适应增强(adaptive boosting, Adaboost)和梯度提升决策树(gradient boosting decision tree, GBDT)等传统集成算法, Xgboost 一方面通过引入正则项和列抽样提高稳定性,另一方面通过并行化策略选择分裂点降低模型的运行时间。将 Xgboost 应用于短期负荷预测输入特征选择,推导过程如下。

将短期负荷预测的数据集 $\mathbf{D} = \{(x(i), y(i)); i = 1, \dots, T, x_i \in \mathbf{R}^n, y_i \in \mathbf{R}\}$, 其中 T 为样本数量, n 为特征数量, 建立回归树模型如下:

$$\hat{y} = \sum_{m=1}^M f_m(x) \quad (1)$$

式中: f_m 表示第 m 棵树, 棵树总数为 M 。以下为目标函数:

$$O = \sum_{i=1}^T l(y(i), \hat{y}(i)) + \sum_{m=1}^M \Omega(f_m) \quad (2)$$

式中: l 为误差函数, 用于计算拟合负荷与实际负荷之间的差距, $\Omega(f_m)$ 表示正则化惩罚项, 表达式如下:

$$\Omega(f_m) = \frac{1}{2} \lambda \sum_{j=1}^T \omega_k^2 + \gamma L \quad (3)$$

式中: ω_k 表示叶子节点 k 的权重, λ 为叶子节点权重对应的惩罚系数, L 代表叶子节点的数量, γ 指代叶子节点数量对应的惩罚系数。

在实际训练中, Xgboost 通过 boosting 机制, 将第 t 次迭代时的回归树模型表示为 $\sum_{m=1}^{t-1} f_m(x)$ 和 $f_t(x)$ 的组合, 因此第 t 次迭代时, 目标函数 O_t 表达式如下:

$$O_t = \sum_{i=1}^T l(y(i), \sum_{m=1}^{t-1} f_m(x(i)) + f_t(i)) + \Omega(f_t) \quad (4)$$

对式(4)进行二阶泰勒展开简化后, 得到第 t 次迭代时 d 的目标函数。

$$O_t = \sum_{i=1}^T \left[\left(\sum_{m=1}^{t-1} g_i \right) \omega_k + \frac{1}{2} \left(\sum_{m=1}^{t-1} h_i + \lambda \right) \omega_k^2 \right] + \gamma L \quad (5)$$

式中: g_i 为损失函数的一阶导数, h_i 为损失函数的二阶导数。叶子节点 j 的最优权重由式(6)计算可得。

$$\omega_k^* = -\frac{G_k}{H_k + \lambda} \quad (6)$$

结合式(5)和(6), 可得最优目标值的计算表达式:

$$O^* = -\frac{1}{2} \sum_{k=1}^L \frac{G_k^2}{H_k + \lambda} + \lambda L \quad (7)$$

式中: G_k 表示全部样本的 g_i 之和, H_k 代表全部样本的 h_i 之和。采用贪心算法进行分支, 遍历当前叶子节点的全部分支条件, 计算分支增益, 表达式如下:

$$Gain = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{G_L + G_R}{H_L + H_R + \lambda} \right] - \gamma \quad (8)$$

式中: L 表示分支后的左叶子, R 表示分支后的右叶子。选择目标函数最小, 分支增益最大的分支条件。

Xgboost 集成训练若干回归树后, 每次分裂获得的增益表明了该分裂特征的重要性, 可以将一个特征的平均增益作为该特征的重要性度量, 平均增益越高, 重要性越强, 计算表达式如下:

$$AGain_j = \frac{1}{FScore_j} \sum_{k=1}^{FScore_j} Gain_k \quad (9)$$

其中, $FScore_j$ 表示特征 j 被选为分裂特征的次数。

1.2 特征优选策略

本文所提输入数据特征选择方法的流程如图 1 所示, 具体步骤如下:

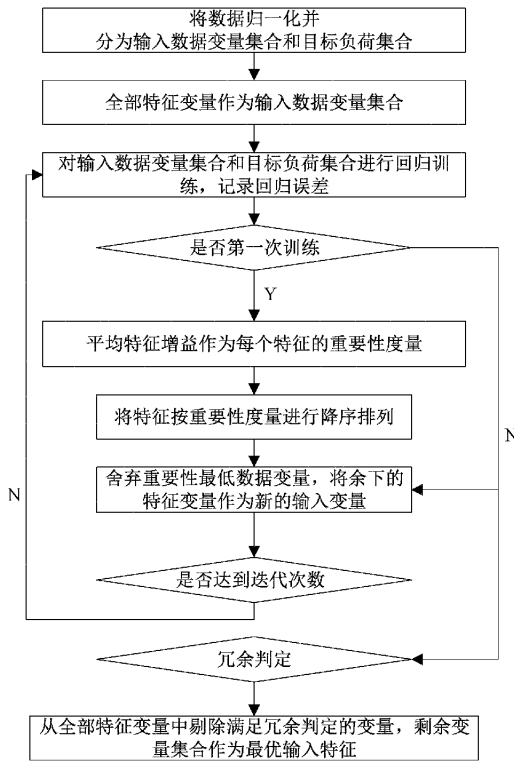


图 1 特征优选策略框架

1) 对数据集 $Z = \{z_1(t), z_2(t), \dots, z_m(t)\}$ 进行归一化处理, 得到归一化后的样本数据集 $\tilde{Z} = \{\tilde{z}_1(t), \tilde{z}_2(t), \dots, \tilde{z}_m(t)\}$, 归一化处理的表达式如式(4)所示。

$$\tilde{z}_j(t) = \frac{z_j(t) - \min_{1 \leq t \leq T} [z_j(t)]}{\max_{1 \leq t \leq T} [z_j(t)] - \min_{1 \leq t \leq T} [z_j(t)]} \quad (10)$$

2) 依据归一化后的样本数据集 $\tilde{Z} = \{\tilde{z}_1(t), \tilde{z}_2(t), \dots, \tilde{z}_m(t)\}$ 生成输入特征集合 $X = \{x(1), x(2), \dots, x(T)\}$ 和对应目标负荷集合 $Y = \{y(1), y(2), \dots, y(T)\}$ 。

3) 通过 Xgboost 对目标负荷 y 和输入特征向量 $x =$

$[x_1, x_2, \dots, x_n]$ 做回归训练, 并记录当前回归误差 e_{cur} 。

4) 将特征平均增益作为每个特征的重要性度量指标 w_j , 按照 w_j 对特征进行降序排列, 舍弃 w_j 最小对应特征, 更新输入特征向量 $x = [x_1, x_2, \dots, x_{n-1}]$, 重复步骤 3)。

5) 对冗余特征进行判定, 当舍弃某一特征 x_j 时, 回归误差 e_{cur} 较未舍弃之前 e_{pre} 一定程度降低则判定该特征为冗余特征。

6) 剔除步骤 5) 中选出的冗余特征, 剩余特征集合作为最优特征集合。

2 DLSTMs-FCN 短期负荷预测模型

DLSTMs-FCN 并联优化的负荷预测模型首先利用 LSTM 提取负荷的时序特性, 然后通过 Dropout 机制优化堆叠的 LSTMs 结构, DLSTMs 在提高模型解释能力的同时防止过拟合, 最后通过 FCN 与 DLSTMs 并联, 补充高分辨率信息。

2.1 LSTM 结构

Hochreiter 等针对梯度爆炸和梯度消失问题提出了 LSTM 结构^[9], 该网络结构通过引入门控机制, 在时序问题中具有更强的挖掘能力。LSTM 网络的单元结构如图 2 所示。

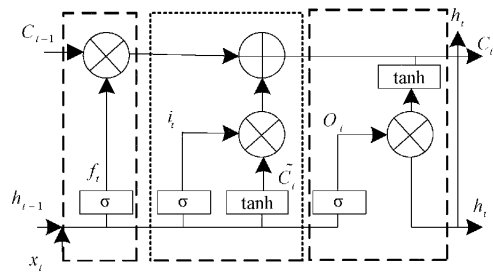


图 2 长短时记忆神经网络单元结构

图 2 中, x_t 表示 t 时刻的输入特征向量; h_t 表示 t 时刻隐层状态值, 对于短期的输入非常敏感; c_t 表示 t 时刻记忆单元状态值, 能够保存长期状态^[9]。长期状态通过输入门(input gate)、输出门(output gate)和遗忘门(forget gate)共同控制, 其中输入门、输出门用来读取、输出和修正参数, 而遗忘门利用 c_{t-1} 选择性遗忘参数。计算公式如下:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \quad (11)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (12)$$

$$g_t = \tanh(W_g[h_{t-1}, x_t] + b_g) \quad (13)$$

$$c_t = f_t c_{t-1} + i_t g_t \quad (14)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (15)$$

$$h_t = o_t \tanh c_t \quad (16)$$

式中: f_t, i_t, o_t 分别代表遗忘门、输入门和输出门; g_t 表示当前输入的细胞状态; c_t 表示当前时刻的细胞状态; W 和 b 分别为对应的权重系数矩阵和偏置项; σ 表示 sigmoid 激活函数, \tanh 表示双曲正切激活函数。

2.2 基于 Dropout 的 DLSTMs 结构

LSTMs 由多个 LSTM 单元在同一时间步长上的堆叠

而成。单层 LSTM 结构深度较浅, LSTMs 通过增加网络深度, 提高模型对时序特征的抽象能力, 以实现更高的预测精度。然而随着网络的加深, LSTMs 可能出现过拟合问题^[8]。

为防止 LSTMs 过拟合, 本文采用 Dropout 算法对 LSTMs 进行改进, 建立 DLSTMs 结构。Dropout 算法是指在训练过程中按照一定的概率设定部分神经元暂时处于失活状态, 降低模型对于局部神经元的依赖, 从而跳出局部最优解^[17]。

定义 $r_{j,t}^{(l)}$ 表示 t 时刻第 l 层神经元 j 的激活状态, 取值为 1 时表示该神经元处于激活状态, 取值为 0 表示该神经元处于非激活状态。 $r_{j,t}^{(l)}$ 服从伯努利分布 (bernoulli), 表达式如下:

$$\Pr\{r_{j,t}^{(l)} = 1\} = p \quad (17)$$

$$\Pr\{r_{j,t}^{(l)} = 0\} = 1 - p \quad (18)$$

$$\tilde{x}_{j,t}^{(l)} = r_{j,t}^{(l)} x_{j,t}^{(l)} \quad (19)$$

式中: $\Pr\{\cdot\}$ 表示 $\{\cdot\}$ 中等式成立的概率, $x_{j,t}^{(l)}$ 表示第 l 层 t 时刻特征 j 的输入, $\tilde{x}_{j,t}^{(l)}$ 表示第 l 层 t 时刻特征 j 的输入以概率 p 随机失活后的输入, 通过输入的随机失活等效神经元的随机失活。基于 Dropout 的 DLSTMs 结构如图 3 所示。

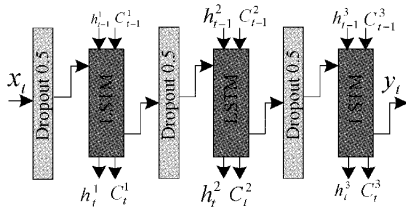


图 3 DLSTMs 结构

DLSTMs 模块首先通过网络的加深提高模型对特征的抽象能力; 然后, 随机忽略部分神经元, 避免 LSTMs 过度依赖某些局部特征, 可在一定程度上降低迭代过程中的过拟合现象。

2.3 基于 FCN 优化的 DLSTMs-FCN 结构

基于 Dropout 的 DLSTMs 虽然解决了过拟合问题, 但是负荷曲线峰谷处的部分局部特征信息可能被忽略。结合感知域较小的 FCN^[19], DLSTMs-FCN 通过并联的结构将提取的抽象特征与负荷波动对齐, 补充高分辨率信息。

1) FCN 结构

FCN 由多层基本卷积块和输出层构成, 输出层采用卷积核大小为 1×1 的卷积输出预测结果。结构如图 4 所示。

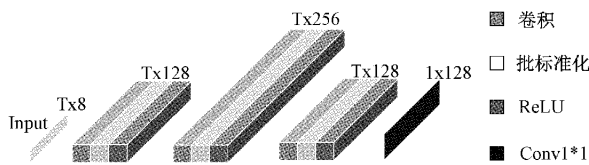


图 4 FCN 结构

基本卷积块包括一维卷积层、批标准化和激活函数。首先, 卷积块通过卷积获取特征信息和特征之间的位置信息; 其次, 采用标准化改变卷积层输出分布, 使得输出数据的分布相对稳定; 最后通过修正线性单元 (linear rectification function, ReLU) 实现非线性化。

一维卷积表达式如下。

$$e_{j,t} = \omega_b + \sum_{t'=1}^d \omega_{j,t'} x_{j,t+t'} \quad (20)$$

式中: $x_{j,t}$ 表示 t 时刻特征 j 的输入, $e_{j,t}$ 表示 t 时刻特征 j 对应卷积层输出, $\omega_{j,t'}$ 表示特征 j 对应张量, ω_b 表示卷积偏置项。卷积核 d 的大小决定卷积网络的感知域范围, 感知域越小, 获得的卷积信息分辨率越高^[20]。

标准化参数表达式如下。

$$u_j = \frac{1}{T} \sum_{t=1}^T e_{j,t} \quad (21)$$

$$\sigma_j^2 = \frac{1}{T} \sum_{t=1}^T (e_{j,t} - u_j)^2 \quad (22)$$

$$\hat{e}_{j,t} = \gamma \frac{e_{j,t} - u_j}{\sqrt{\sigma_j^2 + \epsilon}} + \beta \quad (23)$$

式中: μ_j 表示特征 j 的均值, σ_j^2 表示特征 j 的方差, $\hat{e}_{j,t}$ 表示标准化矩阵, γ 表示标准化方差, β 表示标准化后, ϵ 表示防止除零的极小量。

激活函数 ReLU 表达式如下。

$$y_{j,t} = f(\hat{e}_{j,t}) = \max(0, \hat{e}_{j,t}) \quad (24)$$

2) DLSTMs-FCN 结构

DLSTMs-FCN 模型如图 5 所示。负荷特征输入到该模型后, 首先通过 DLSTMs 抽象负荷时序特性, 得到高维信息, 再使用 FCN 获取负荷特征的高分辨率信息, 最后通过全连接网络 (fully connected network, FNN) 预测目标负荷。

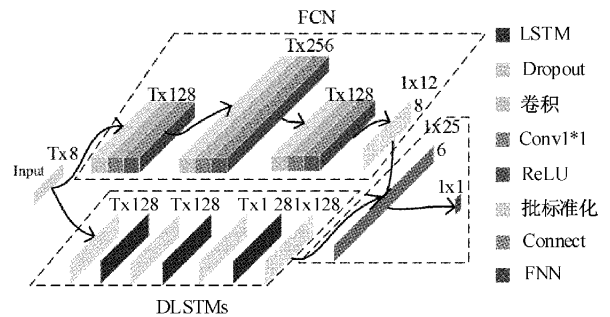


图 5 DLSTMs-FCN 结构

3 算例分析

本文选取澳大利亚新南威尔士州 2006 年~2010 年间 3 652 天的 87 648 条电力负荷及相关天气数据作为数据集^[21]。该数据集以 0.5 h 为采样间隔, 记录负荷、电价、天气等数据, 具有采样频率高、数据量大以及特征丰富等优点。

首先使用本文所提特征优选方法求解最优特征集合；其次将本文所提特征优选方法与基于 Pearson 的特征选择方法,基于 MIC 的特征选择方法和基于 Xgboost 的直接特征筛选算法进行对比,以验证本文所提特征优选方法的有效性;最后将本文所提方法与 LSTM、LSTMs、DLSTMs、ALSTMs (Attention-LSTMs)和 CNN-LSTMs 共同进行实验。其中 LSTM、LSTMs 为传统 LSTM 算法作为对比项,验证 Dropout 在 LSTMs 中应用的有效性,而 Attention-LSTMs 和 CNN-LSTMs 为其他文献中最新的改进算法与本文所提改进进行对比,充分验证本文所提方法性能。

3.1 评价指标

选取均方根误差(root mean square error, RMSE)评价负荷预测误差。

$$e_{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - y'_i)^2} \quad (25)$$

式中: n 为预测次数, y'_i 为第 i 次负荷预测值, y_i 为第 i 次负荷实际值。

选取误差的标准差 (standard deviation of error, SDE) σ_e 评价模型预测误差的波动程度,检验模型预测结果的稳定性。

$$e_i = y_i - y'_i \quad (26)$$

$$\bar{e} = \frac{1}{n} \sum_{i=1}^n (e_i) \quad (27)$$

$$\sigma_e = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (e_i - \bar{e})^2} \quad (28)$$

式中: e_i 为第 i 次的负荷预测误差, \bar{e} 为 n 次误差平均值。

3.2 特征优选

在文献[9]选择特征的基础上,补充日、月、周、时和节假日 5 种特征并进行回归训练。特征和对应权重如表 1 所示。

表 1 重要性度量排序

特征	特征描述	重要性度量	排序	RMSE
l^{t-1}	前一时刻的负荷值	6.900 144	1	246.2
l^{t-2}	前二时刻的负荷值	0.053 131	2	128.8
t^{hou}	0.5~24 代表一天的 48 个时刻	0.046 920	3	94.97
p	电价	0.036 468	4	96.31
l^{t-48}	前一日对应时刻的负荷值	0.030 179	5	93.85
h	温度/°C	0.013 216	6	83.05
l^{t-49}	前一日对应前一时刻的负荷值	0.010 249	7	75.58
t^{wee}	1~7 代表周一~周日	0.007 825	8	67.11
t^{mou}	1~12 代表 1 月~12 月	0.003 087	9	68.63
c	湿度/%RH	0.001 568	10	69.88
t^{hol}	0 代表非节假日,1 代表节假日	0.000 909	11	67.21
t^{day}	1~31 代表 1 号~31 号	0.000 769	12	68.37

将表 1 中 12 种特征通过权重的数值大小进行排序,得到重要性排序。按照排序结果舍弃低重要性特征,对余下特征进行回归训练,得到的均方根误差如表 1 中 RMSE 所示。可得当舍弃特征 c 、 t^{mou} 和 t^{day} 时预测误差降低,判定 t^{day} 、 c 和 t^{mou} 为冗余特征。

对冗余特征进行可解释性分析:特征 t^{day} 和 t^{mou} 为日信息和月信息,与 t^{wee} 和 t^{hol} 有效信息重复度较高;特征 c 为该时刻湿度对负荷影响较小,无效信息较多,并且 t^{day} 、 c 和 t^{mou} 的重要性度量较低。

因此本文舍弃冗余特征 9,10,12,选择 $[l^{t-1}, l^{t-2}, l^{t-48}, l^{t-49}, p, t^{hou}, t^{wee}, t^{hol}, h]$ 为输入特征组合。

3.3 不同特征选择算法效果对比

本文所提特征优选算法得到的特征优选集合为 $[l^{t-1}, l^{t-2}, l^{t-48}, l^{t-49}, p, t^{hou}, t^{wee}, t^{hol}, h]$; 基于 Xgboost 的直接特征筛选算法选取重要性度量大于 0.01 的特征,得到的特征集合为 $[l^{t-1}, l^{t-2}, t^{hou}, p, l^{t-48}, h, l^{t-49}]$; 基于 MIC 特征

选择算法选取相关性大于 0.5 的特征^[9],得到特征集合为 $[l^{t-1}, l^{t-2}, l^{t-48}, l^{t-49}, p, h, c, t^{hou}]$; 基于 Pearson 特征选择算法选取相关性大于 0.5 的特征,得到的特征集合为 $[l^{t-1}, l^{t-2}, l^{t-48}, l^{t-49}, t^{hou}]$ 。分别将不同特征选择算法选择的输入特征集合作为输入特征,对负荷进行预测,计算对应的 RMSE 和 SDE,结果分别如图 6 和 7 所示。

由图 6 和 7 可知,不同特征选择算法得到的特征集合作进行负荷预测的 RMSE 和 SDE 存在一定差异,全特征变量集合 RMSE 最大,并且 SDE 也最大;基于 Pearson 特征筛选方法得到的特征集合中只包含了历史负荷和时信息,提供的信息量较少;基于 MIC 特征选择方法和基于 Xgboost 的直接特征筛选算法均考虑了电价和气候特征与负荷之间的相关性,相较于前两者的预测误差和误差波动有一定减少,但是采用阈值进行筛选,缺少对多特征综合影响的处理。本文基本 Xgboost 的特征优选算法考虑了多特征的共同影响,能够有效去除重要性低、信息冗余的特征,相较

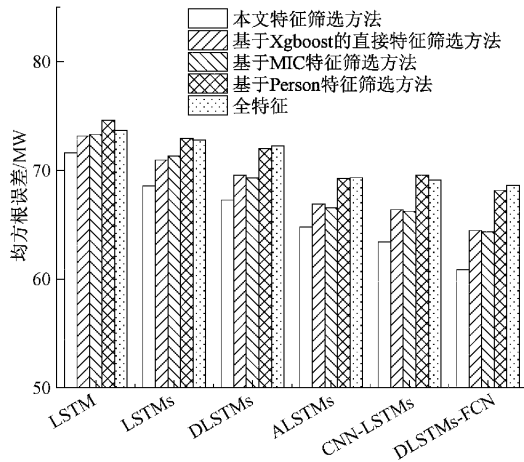


图 6 不同特征选择算法的 RMSE 对比

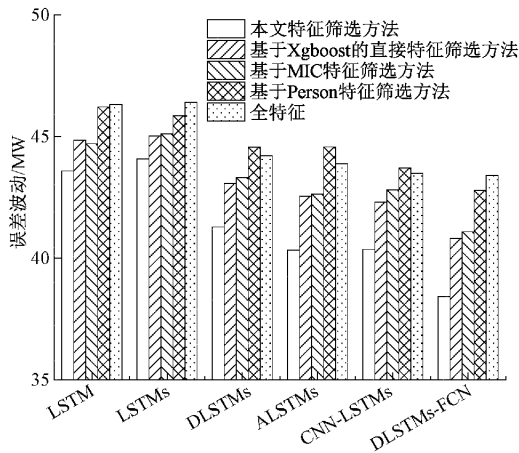


图 7 不同特征选择算法的 SDE 对比

于 MIC, Person 特征筛选算法,不同的模型 RMSE 降低了 1.3%~4%,SDE 降低了 4%~10%,因此本文特征优选策略得到特征集合更有利于以 LSTM 为基础的模型进行负荷预测。

3.4 不同 LSTM 负荷预测模型对比

将 LSTM, LSTMs, DLSTMs, Attention-LSTMs, CNN-LSTMs 和本文所提模型 DLSTMs-FCN 进行短期负荷预测实验,以验证本文所提优化模型的准确性和稳定性。其中,LSTM 的隐藏层统一设置为 128, LSTMs 采用三层隐藏层,CNN-LSTMs 为 LSTMs 与 CNN 串联,其中 CNN 的卷积核大小为 5,步长为 1^[10]。

统一输入特征组合为本文特征优选集合 $[l^{t-1}, l^{t-2}, l^{t-18}, l^{t-19}, p, t^{hour}, t^{week}, t^{hol}, h]$, 模型训练样本为 2006 年 1 月~2008 年 12 月的数据信息,验证样本为 2009 年 1 月~12 月的数据信息,测试样本为 2010 年 1 月~12 月数据信息,每个样本的时间步长为 48。每个模型的预测评价指标如图 8 所示。

由图 8 可知,本文所提优化模型 DLSTMs-FCN 的 RMSE 和 SDE 较其他模型最优。相对于单层 LSTM 模

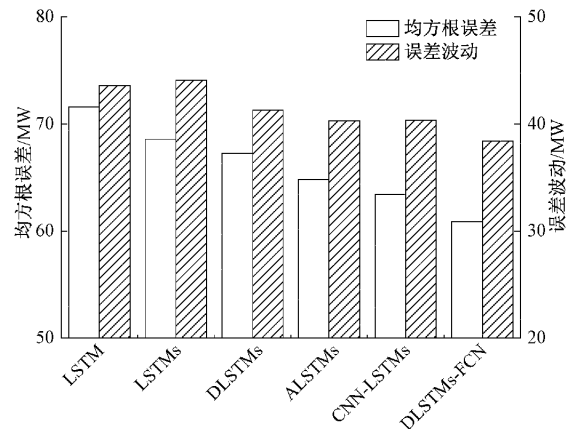


图 8 不同 LSTM 负荷预测模型对比

型,LSTMs 的 RMSE 降低了 4.2%,但是 SDE 却增加了 1.1%。说明 LSTMs 的模型抽象能力较高,整体预测效果有一定的优化,但是模型过拟合导致其预测稳定性较低。相较于 LSTMs,DLSTMs 的 RMSE 下降了 1.89%同时 SDE 下降了 6.3%,预测稳定性有较大提升,在保障模型抽象能力的同时避免过拟合。将本文所提优化模型 DLSTMs-FCN 与文献[10]中提出的 ALSTMs 和文献[11]中提出的 CNN-LSTMs 相比,DLSTMs-FCN 的 RMSE 分别降低了 6%和 4%,SDE 分别降低了 4.7%和 4.8%,预测精度和稳定性均有明显提升。这是因为 FCN 与 LSTM 模型特性存在较大差异性,为 DLSTMs 增加并联的 FCN 结构,可以有效补充高分辨率信息,缓解 LSTM 面对长时序列下的重要信息丢失问题,提升模型整体的信息提取能力。

图 9 表示 2010 年 1 月 1 日 0 点~24 点的预测结果图,其中灰色方块放大图分别表示 3 点~6 点的负荷波谷,4 点~7 点的负荷波峰。该日为澳大利亚的元旦和圣诞节,同时处于澳大利亚夏季,盛行西北季风,且为台风期,社会和环境因素相对复杂。

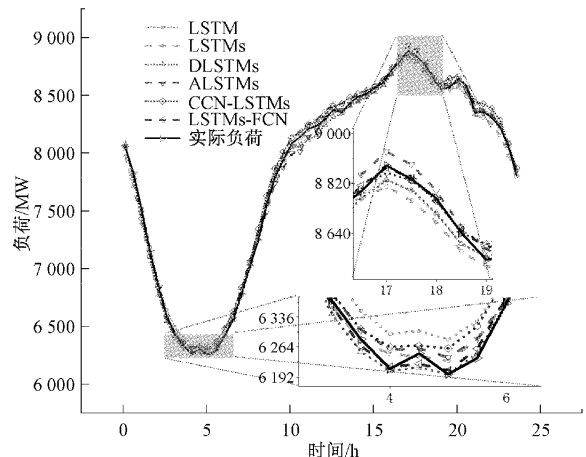


图 9 2010 年 1 月 1 日负荷预测结果

从整体波形观察,本文所提优化模型 DLSTMs-FCN 得到的预测负荷与实际负荷曲线基本拟合。观察波峰和

波谷细节,可以发现相较于其他几种模型,DLSTMs-FCN 的预测结果对于负荷高峰和负荷低谷的处理更好。针对该日实际负荷在波谷处 4:00~5:00 时段出现的跳变,DLSTMs-FCN 所得预测结果虽然与实际存在一定差距,但仍优于其余模型。

综上,本文采用基于 Xgboost 的特征优选算法获取更有利于短期负荷预测的特征集合,并通过与 FCN 并联的方式为 DLSTMs 补充高分辨率信息,提高负荷高峰和负荷低谷处理能力,将 RMSE 降低至 60.863 MW,SDE 降低至 38.41 MW,相较于 ALSTMs 和 CNN-LSTMs 模型,预测精度更高,预测稳定性更好。

4 结 论

本文提出了特征优选策略和 DLSTMs-FCN 并联优化结构的短期负荷预测方法,并且在公开数据集上与多种特征筛选方法和优化 LSTM 模型进行了实验对比,验证本文的短期负荷预测模型的有效性和可靠性。

本文所提特征优选算法可以更有效的排除冗余特征,得到特征集合更有利于以 LSTM 为基础的模型进行负荷预测。

本文提出的 DLSTMs-FCN 并联网络结构在短期负荷预测中精度更高和稳定性更强。

参考文献

- [1] 王晓辉,邓威威,齐旺. 基于超参数优化的短期电力负荷预测模型[J]. 国外电子测量技术, 2022, 41(6): 152-158.
- [2] 杨邓,杨俊杰,胡晨阳,等. 基于改进 LSSVM 的短期电力负荷预测[J]. 电子测量技术, 2021, 44(18): 47-53.
- [3] DUDEK G, PELKA P, SMYL S, et al. A hybrid residual dilated LSTM and exponential smoothing model for midterm electric load forecasting[J]. IEEE Transactions on Neural Networks and Learning Systems, 2021, DOI:10.1109/TNNLS.2020.3046629.
- [4] CHODAKOWSKA E, NAZARKO J, NAZRKO Ł, et al. ARIMA models in electrical load forecasting and their robustness to noise[J]. Energies, 2021, 14(23): 1-2.
- [5] 乔石,王磊,张鹏超,等. 基于模态分解及注意力机制长短时间网络的短期负荷预测[J]. 电网技术, 2022, 1(13): 5-14.
- [6] YUAN T, JIANG D S, HUANG S Y, et al. Recurrent neural network based short-term load forecast with spline bases and real-time adaptation[J]. Applied Sciences, 2021, 11(13): 9-20.
- [7] CIECHULSKI T, OSOWSKI S. High precision LSTM model for short-time load forecasting in power systems[J]. Energies, 2021, 14(11): 19-29.
- [8] 高超,孙谊嫒,赵洪峰,等. 基于 ICOA-LSTM 的短期负荷预测研究[J]. 电子测量技术, 2022, 45(13): 88-95.
- [9] IBRAHIM B, RABELO L. A deep learning approach for peak load forecasting: A case study on panama[J]. Energies, 2021, 14(11): 1-2.
- [10] 彭文,王金睿,尹山青,等. 电力市场中基于 Attention-LSTM 的短期负荷预测模型[J]. 电网技术, 2019, 43(5): 1745-1751.
- [11] 魏健,赵红涛,刘敦楠,等. 基于注意力机制的 CNN-LSTM 短期电力负荷预测方法[J]. 华北电力大学学报(自然科学版), 2021, 48(1): 42-47.
- [12] SHANG C, GAO J W, LIU H B, et al. Short-term load forecasting based on PSO-KFCM daily load curve clustering and CNN-LSTM model [J]. IEEE ACCESS, 2021, 9(9): 2169-3536.
- [13] 夏飞,张洁,张浩,等. 基于 BIC 准则和加权皮尔逊距离的居民负荷模式精细识别及预测[J]. 电子测量与仪器学报, 2020, 34(11): 33-42.
- [14] TONG X, WANG J Y, ZHANG C L, et al. LS-LSTM-AE: Power load forecasting via Long-Short series features and LSTM-Autoencoder[J]. Energy Reports, 2022, 8(S1): 2352-4847.
- [15] 方明宽,宁静,陈春俊. 高速列车横向蛇行失稳的 EEMD-CNN-LSTM 预测方法[J]. 中国测试, 2021, 47(7): 79-83.
- [16] 郑睿程,顾洁,金之俭,等. 数据驱动与预测误差驱动融合的短期负荷预测输入变量选择方法研究[J]. 中国电机工程学报, 2020, 40(2): 487-500.
- [17] 车昱娇,陈云霞,崔宇轩. KPCA 和改进 LSTM 在滚动轴承剩余寿命预测中的应用研究[J]. 电子测量与仪器学报, 2021, 35(2): 109-114.
- [18] FAZLE K, SOMSHUBRA M, HOUSHANG D, et al. Multivariate LSTM-FCNs for time series classification[J]. Neural Networks, 2019, 116(1): 1-9.
- [19] 石欣,田文彬,冷正立,等. 基于 CFD 和 LightGBM 算法的建筑室内温度全局预测模型[J]. 仪器仪表学报, 2021, 42(1): 237-247.
- [20] 杜艳玲,崔建华,魏泉苗,等. 基于改进 FCN 的多极化 SAR 影像海上溢油检测[J]. 激光与光电子学进展, 2022, 59(4): 321-328.
- [21] Australian Energy Market Operator. AEMO's first national gas forecasting report[R/OL]. <http://www.aemo.com.au>, [2014-11-17].

作者简介

孟金鑫(通信作者),硕士研究生,主要研究方向为智能数据挖掘,电力系统负荷预测。

E-mail: scu_mjx@163.com

黄山,教授,博士生导师,主要研究方向为图形图像,智能交通。

E-mail: scu_hs@163.com

印月,工程师,主要研究方向为电力系统设计。

E-mail: yyinn2016@126.com