

基于改进 Bi-LSTM 和 XGBoost 的电力负荷组合预测方法

代业明, 周琼

(青岛大学 商学院, 青岛 266071)

摘要: 电力负荷预测在平衡能源分配、经济性和电力系统安全可靠运行方面发挥着重要作用, 精准的负荷预测可以降低电力运行的成本和风险, 提高电网环境效益和经济效益。首先根据加权灰色关联投影算法对数据进行预处理, 然后应用注意力 (Attention) 机制来改进双向长短期记忆 (Bi-LSTM) 模型, 并结合极端梯度提升 (XGBoost) 模型构建一种由误差倒数法确定权重的电力负荷组合预测模型, 从而得到一种新的短期电力负荷预测方法。通过新加坡电力市场数据集对该方法进行评估, 结果显示, 该方法的预测结果比单一预测方法更加接近真实数据且误差更小, 具备有效性、精准性和实用性的优势。

关键词: 短期负荷预测; 双向长短期记忆网络; 极端梯度提升; 误差倒数组合法; 组合模型
中图分类号: TM 73 **文献标志码:** A

Power load combination forecasting method based on improved Bi-LSTM and XGBoost

DAI Yeming, ZHOU Qiong

(School of Business, Qingdao University, Qingdao 266071, China)

Abstract: Power load forecasting plays an important role in balancing energy distribution, economy, safe and reliable operation of power system. Accurate load forecasting can reduce the cost and risk of power operation and improve the environmental and economic revenues of power grid. Firstly, the weighted grey relational projection algorithm was used to process the data, then the attention mechanism was applied to improve the bi-directional long short-term memory (Bi-LSTM) model. By combining with eXtreme gradient boosting (XGBoost) model, a power load combination forecasting model with the weight determined by the error reciprocal method was developed, so as to obtain a new short-term power load forecasting method. Finally, based on the Singapore power market data set, the results of the proposed prediction method show that our prediction results are closer to the real data and have less error than the single prediction methods, which reflect the effectiveness, accuracy and practicability of proposed method.

Keywords: *short-term load forecasting; bi-directional long short-term memory (Bi-LSTM); eXtreme gradient boosting (XGBoost); error reciprocal method; combined model*

随着电力需求日益增加,传统电网在集中配电、人工监控和恢复、双向通信等方面开始遇到挑战^[1-2],智能电网的出现为解决上述困难提供技术可能,有助于监控电力生产、传输和消耗,并平衡三者关系^[3]。但由于气候、经济、环境等不确定因素的影响,电力负荷波动较大^[4],难以简单对其进行预测。为了确保电力系统安全运行,减少电力消耗,满足市场需求,平衡电力负荷的供需关系,最大化经济效益,对电力负荷进行精准预测十分必要。

现有电力负荷预测方法主要分为4类:传统预测、现代预测、混合预测和组合预测。其中,传统预测方法包括时间序列分析^[5]、回归分析^[6]和其他统计方法,能较处理好简单线性问题和估计未来电力负荷,但在处理非线性问题时效果不佳。于是,基于非线性映射的现代预测技术逐渐用来预测非线性问题,主要包括模糊逻辑、灰色系统、人工神经网络、支持向量回归。但这些方法依然存在计算复杂^[7-8]、泛化能力差^[9]和过拟合^[10]等固有局限,对电力负荷精准预测带来新的挑战。

针对上述不足,通过粒子群^[11]、贝叶斯^[12]等参数优化算法以及对数据进行预处理,提出了混合预测模型。从参数优化角度,Wang等^[13]使用混合支持向量回归方法预测中长期负荷,并使用基于嵌套策略和状态转移算法的分层方法优化预测模型的参数;Barman等^[14]使用灰狼优化器优化支持向量机参数并预测了受文化或宗教仪式等社会因素显著影响的电力需求。通过数据分解^[15]和特征选择技术^[16]对数据进行预处理,有助于消除异常值,纠正数据错误和提高数据质量。因此,混合预测方法比前两种预测方法表现更佳。

为进一步改进和优化预测模型,克服传统、现代和混合预测方法中单一预测模型的固有缺陷,将两个以上的不同预测模型以特定加权方法组合,从而得到组合预测方法。如Xiao等^[17]采用布谷鸟搜索算法来优化组合模型的权重系数,实验结果表明所提出的组合预测模型可以产生更低的误差,能够提供稳定、高精度的负荷预测;林锦顺等^[18]、邓带雨等^[19]和Chu等^[20]也证明了组合预测模型的性能优于单一模型;陈振宇等^[21]使用

长短期记忆网络(long short-term memory, LSTM)和极端梯度提升(eXtreme gradient boosting, XGBoost)模型预测电力负荷,然后根据误差倒数法为两个模型分配权重,从而尽可能减小误差和提高预测精度;庄家懿等^[22]结合MAPE-RW算法进行模型组合初始权重设置,对最佳权重进行搜索,构建了CNN-LSTM-XGBoost组合预测模型,比单一预测模型误差显著降低;Nie等^[23]则使用多目标灰狼算法(multi-objective grey wolf algorithm, MOGWO)分别对径向基函数网络(radial basis function network, RBF)、广义回归神经网络(generalized regression neural network, GRNN)和极限学习机(extreme learning machine, ELM)赋予权重,建立了一种基于群智能优化的组合预测模型,有效降低了单一模型适应性弱的不利影响,使准确性和适应性得到迅速提高。

从以上文献可知,组合预测方法对预测的改进不仅需要多个预测模型,还需要计算每个模型所占权重,并没有强调数据预处理的重要性,并且通常组合现有成熟预测模型。在此背景下,本文首先使用加权灰色关联投影算法(weighted grey relational projection, WGRP)对数据进行预处理,消除节假日的不利影响;其次,在被广泛认为是一种能够充分隐藏信息并获得良好预测效果的Bi-LSTM预测模型基础上,将具有大规模并行处理、分布式信息存储、可接受的自组织和自学习能力等优点的注意力机制添加到Bi-LSTM模型中,消除不合理的影响,并强调关键输入数据影响,得到了“Attention-Bi-LSTM”模型;最后,为了避免单一预测模型缺陷,再将通过添加正则项来控制模型复杂性和防止过度拟合、提高模型泛化能力的XGBoost^[21]模型与Attention-Bi-LSTM模型组合,得到“Attention-Bi-LSTM + XGBoost”电力负荷组合预测模型,并运用新加坡电力市场数据进行评估。

1 基本理论

1.1 加权灰色关联投影算法

加权灰色关联投影算法^[16]是一种将各因素之间发展趋势的相似或相异程度,即“灰色关联度”作为衡量各因素关联程度的一种方法。该方

法不受样本量多少的限制,可以全面分析指标间相互关系,反映各决策方案与理想方案之间接近程度,并对有无规律的样本量都同样适用,结果具有一般性、计算量小的特点。

首先,选择 n_1 个样本的数据和待预测样本的数据,计算它们之间的关系系数,并构造以下灰色关联矩阵。

$$A = \begin{bmatrix} A_{01} & \cdots & A_{0m_1} \\ \vdots & & \vdots \\ A_{n_1 1} & \cdots & A_{n_1 m_1} \end{bmatrix} \quad (1)$$

式中, $A_{n_1 m_1}$ 表示第 n_1 个样本中第 m_1 个因子的灰色关联系数。

然后,使用熵权法计算各影响因素的权重,并对灰色关联矩阵进行加权,得到加权灰色关联矩阵,如下所示:

$$A' = A\gamma^T \begin{bmatrix} \gamma_1 & \cdots & \gamma_{m_1} \\ \vdots & & \vdots \\ \gamma_1 A_{n_1 1} & \cdots & \gamma_{m_1} A_{n_1 m_1} \end{bmatrix} \quad (2)$$

式中: γ 表示各影响因素的权重;矩阵中的第一行表示为待预测样本的行向量 A'_0 ;其他历史样本的行向量表示为 A'_i ;每个 A'_0 和 A'_i 之间的夹角是样本的灰色投影角,用 θ_i 表示,并计算出 $\cos\theta_i$ 为

$$\cos\theta_i = \frac{\sum_{j=1}^{m_1} \gamma_j A_{ij} \gamma_j}{\sqrt{\sum_{j=1}^{m_1} (\gamma_j A_{ij})^2} \sqrt{\sum_{j=1}^{m_1} \gamma_j^2}} \quad (3)$$

因此,每个历史样本和待预测样本之间的加权灰关联投影值 B_i 为

$$B_i = \frac{\sum_{j=1}^{m_1} \gamma_j A_{ij} \gamma_j}{\sqrt{\sum_{j=1}^{m_1} \gamma_j^2}} \quad (4)$$

最后,将得到的投影值从大到小排序,选择投影值较大的样本作为相似样本进行替换。

1.2 Attention-Bi-LSTM 电力负荷预测模型

1.2.1 Bi-LSTM 模型

为了改善循环神经网络中的梯度爆炸和消失的问题,LSTM用门控机制来控制信息的更新或丢弃,引入了输入门、遗忘门、输出门^[24],以此去除一些对当下情况不重要的内容,使得信息保存时间延长,可以保存一些时间较为久远的信息。LSTM门的输入均为当前时间步输入 X_t 与上一时间步隐藏状态 H_{t-1} ,输出由激活函数 sigmoid 函数

(σ)的全连接层计算得到。其总体框架为

$$\text{输入门 } I_t = \sigma(X_t W_{xi} + H_{t-1} W_{hi} + b_i) \quad (5)$$

$$\text{遗忘门 } F_t = \sigma(X_t W_{xf} + H_{t-1} W_{hf} + b_f) \quad (6)$$

$$\text{门控单元 } \tilde{C}_t = \tanh(X_t W_{xc} + H_{t-1} W_{hc} + b_c) \quad (7)$$

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \quad (8)$$

$$\text{输出门 } O_t = \sigma(X_t W_{xo} + H_{t-1} W_{ho} + b_o) \quad (9)$$

式中: X_t 为给定时间步 t 的小批量输入向量; W_{xi} 和 W_{hi} 为输入门的权重矩阵; b_i 为输入门的偏置项; W_{xf} 和 W_{hf} 为遗忘门的权重矩阵; b_f 为遗忘门的偏置项; $\tilde{C}_t$ 为短期记忆需要计算的候选记忆细胞; W_{xc} 和 W_{hc} 为门控单元的权重矩阵; b_c 为门控单元的偏置项; C_t 为当前时刻门控单元的状态, C_{t-1} 表示上一次的单元状态; W_{xo} 和 W_{ho} 为输出门的权重矩阵, b_o 为输出门的偏置项。使用值域在 $[-1,1]$ 的 $\tanh$ 函数作为激活函数,通过按元素乘法 $\odot$ 控制隐藏状态中的信息流动。

输出门 O_t 控制记忆细胞到隐藏状态的信息流动,得到最终输出 H_t 为

$$H_t = O_t \odot \tanh C_t \quad (10)$$

与 LSTM 不同,Bi-LSTM 方法由前向 LSTM 和后向 LSTM 组合而成,在对数据特征提取的时候会考虑到隐藏在数据中的整体信息,通过正向和反向两个角度来进行特征提取,然后将双向提取的结果按照特定方式进行结合。从两个维度进行总结,在一定程度上消除单一 LSTM 中输入数据顺序对最终结果的不利影响,使结果更全面。

1.2.2 Attention 机制

目前,Attention 机制在计算机视觉、手写识别等领域已被广泛使用。应用于深度学习中的 Attention 机制可以理解为如何从输入数据中过滤出关键信息,并对这些关键信息赋予更高的权重以作出有效决策。为了强调输入数据对输出数据的不同影响力,优化数据的特征提取,提高预测性能,本文在 Bi-LSTM 的基础上引入 Attention 机制,可以计算出注意力的概率分布,消除输入数据对输出数据不合理的影响,从而提高关键输入数据的影响力。其结构如图 1 所示,相关计算见文献[25]。

1.3 XGBoost 电力负荷预测模型

XGBoost 通过不断添加树,不断进行特征分裂来生长一棵树,每添加一个树,就是学习一个新函数,去拟合上次预测的残差。当训练完成得到 k 棵树,每棵树中包含若干叶子结点,每个叶子

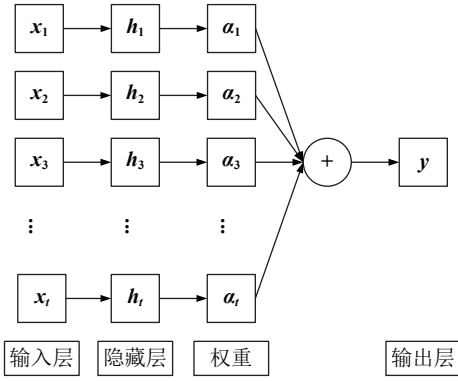


图1 注意力机制结构

Fig. 1 Structure of Attention mechanism

结点就对应一个分数。最后只需要将每棵树对应的分数加起来就是该样本的预测值, 即各样本和它们对应的权值乘积累加^[21]。因此, 随着迭代次数的增加, 预测精度也不断提高。XGBoost 模型如下:

$$\hat{y}_i = \sum_{j=1}^n w_j x_{ij} \quad (11)$$

式中: $\hat{y}_i$ 为预测值; n 为树的数量; w_j 为权值; x_{ij} 为样本数据,

每一次迭代, 都在现有树的基础上, 增加一棵树去拟合前面树的预测结果与真实值之间的残差, 其迭代过程如下:

$$\begin{aligned} \hat{y}_i^{(0)} &= 0 \\ \hat{y}_i^{(1)} &= f_1(x_i) = \hat{y}_i^{(0)} + f_1(x_i) \\ \hat{y}_i^{(2)} &= f_1(x_i) + f_2(x_i) = \hat{y}_i^{(1)} + f_2(x_i) \\ &\vdots \end{aligned}$$

$$\hat{y}_i^{(t)} = \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i) \quad (12)$$

式中: $\hat{y}_i^{(t)}$ 为训练 t 轮后的模型; $\hat{y}_i^{(t-1)}$ 为保留的上一轮添加的函数; $f_t(x_i)$ 为新加的函数。

XGBoost 的目标函数如下所示:

$$Obj^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t)}) + \sum_{i=1}^t \Omega(f_i) = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (13)$$

$$\Omega f(t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \quad (14)$$

最终目标是找到使目标函数最小的 f_t , 式(13)中的 $\sum_{i=1}^n \Omega(f_i)$ 为正则项, 决定了树的复杂程度, 值越小, 复杂度越低, 泛化能力越强。式(14)中:

T 为叶子结点的个数; ω 为叶子结点的分数; γ 的用途为控制叶子结点的个数; λ 保证叶子结点的分数不至于太大, 有效防止过拟合。

为找到一个 f_t 使目标函数最小, 在 $f_t = 0$ 处进行泰勒二阶展开, 得到的目标函数近似为

$$\tau^{(t)} \approx \sum_{i=1}^n \left[l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \quad (15)$$

式中: $g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$ 为一阶导数; $h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$ 为二阶导数。

由于前 $t-1$ 棵树的预测分数和 y 的残差不会对目标函数的优化产生影响, 因此直接去掉, 将目标函数继续简化为

$$\tilde{\tau}^{(t)} = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) \quad (16)$$

式(16)将每个样本的损失函数值叠加起来, 将所有以同一个叶子结点的样本进行重组, 过程如下:

$$\begin{aligned} Obj^{(t)} &\approx \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t) = \\ &\sum_{i=1}^n \left[g_i \omega_{q(x_i)} + \frac{1}{2} h_i \omega_{q(x_i)}^2 \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 = \\ &\sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) \omega_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) \omega_j^2 \right] + \gamma T \end{aligned} \quad (17)$$

根据改写后得到的叶子结点分数 ω 的一元二次函数, 可求得最优的 ω_j^* 和目标函数为

$$\omega_j^* = - \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \quad (18)$$

$$Obj = - \frac{1}{2} \sum_{j=1}^T \frac{\left(\sum_{i \in I_j} g_i \right)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \quad (19)$$

为了方便计算, 符合数据输入的要求, 数据要预先被归一化处理, 具体方法如式(20)所示, 使数据被限制在 $[0, 1]$ 的范围内。

$$x^n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (20)$$

式中: $x_{\min}$ 为电力负荷数据中的最小值; $x_{\max}$ 为电力负荷数据中的最大值; x 为需要归一化处理的数据。

2 Attention-Bi-LSTM+XGBoost 电力负荷组合预测模型

2.1 权重赋予方式

在误差倒数法中，误差较小的预测模型被赋予较大权重，于是组合预测模型的总体误差可以显著降低。为确定权重系数，本文采用误差倒数法为模型分配权重，公式如下：

$$f_t = \omega_1 f_{1t} + \omega_2 f_{2t}, t = 1, 2, \cdots, n \tag{21}$$

$$\omega_1 = \frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2} \tag{22}$$

$$\omega_2 = \frac{\varepsilon_1}{\varepsilon_1 + \varepsilon_2} \tag{23}$$

式中： ω_i 为权重值； f_{it} 为 Attention-Bi-LSTM 和 XGBoost 所得出的预测值； $\varepsilon_1, \varepsilon_2$ 分别为预测模型

2.2 Attention-Bi-LSTM+XGBoost 电力负荷组合预测模型

与现有预测模型不同，Attention-Bi-LSTM 模型不仅可以充分考虑隐藏在输入数据中的整体信息，而且可以从两个维度获得更全面的结果，同时也强调了关键输入数据的影响。因此，使用 Attention-Bi-LSTM 模型可以提高预测结果的准确性；此外，XGBoost 作为一种新兴的预测模型，具有较低复杂度，可以防止过度拟合，并具有良好的预测性能。本文首先使用 Attention-Bi-LSTM 和 XGBoost 方法预测电力负荷，并获得相应的误差；然后根据误差结果，采用误差倒数法计算上述两种模型的权重；最终将上述两种模型的不同预测结果组合起来，从而克服单一预测模型各种固有缺陷。该预测模型框架如图 2 所示。

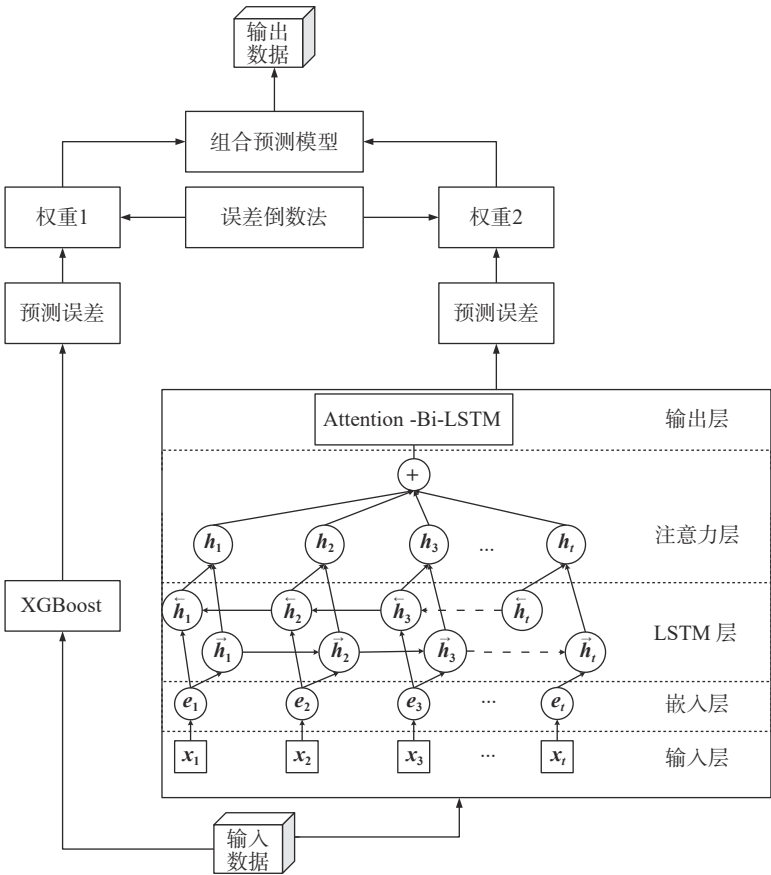


图 2 Attention-Bi-LSTM + XGBoost 组合预测模型框架

Fig.2 Framework of Attention-Bi-LSTM + XGBoost combined prediction model

3 Attention-Bi-LSTM+XGBoost 电力负荷组合预测方法

本文所提出的组合预测方法预测过程如图 3

所示，预测步骤主要分为以下 4 个阶段：
第一阶段，预处理数据。首先，本文选择了几个影响最大的特征，如时间、日类型、假日类型、实时价格，然后利用 WGRP 算法对节假日数

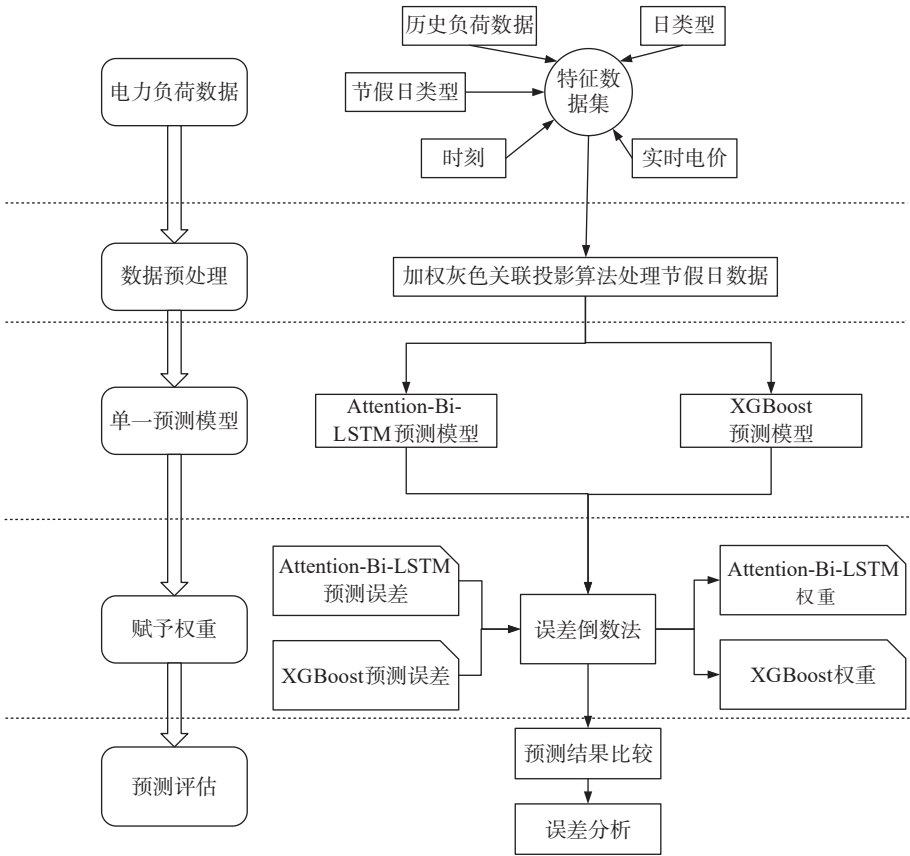


图 3 组合预测模型框架图

Fig.3 Frame diagram of combined forecasting model

据进行处理，使数据更具一般性，最后将数据标准化。

第二阶段，单一模型预测。使用注意力机制对 LSTM 模型进行优化，消除不合理因素的影响，突出关键输入数据的影响，使结果更加全面。Attention-Bi-LSTM 和 XGBoost 模型都用于预测同一数据集，并根据预测结果为两个模型组合作准备。

第三阶段，权重赋予。在用 Bi-LSTM 和 XGBoost 方法对电力负荷数据进行预测后，根据预测的误差使用误差倒数法获得权重，并对单个模型进行加权，形成了 Attention-Bi-LSTM + XGBoost 组合预测模型。

第四阶段，预测评估。通过比较 6 个基准模型和新加坡电力市场实际数据的预测误差，评估该方法能否提高电力负荷预测的准确性。

4 模型评价

本节基于新加坡电力市场数据集评估和讨论所提出的预测方法性能。由于 LSTM，Bi-LSTM，

Attention-Bi-LSTM， Attention-LSTM， Attention-RNN 和 XGBoost 与本文的模型具有很高的相关性，因此作为基准模型与本文提出的方法进行比较。

4.1 数据集和实验环境

本文选取新加坡电力市场 2019 年 1 月 1 日至 2021 年 1 月 7 日的数据，并考虑了日类型、时刻、节假日类型、实时电价等因素，历史电力负荷数据的采样周期为 1 h，2019 年 1 月 1 日至 2020 年 12 月 31 日的数据为训练数据，2021 年 1 月 1 日至 2021 年 1 月 7 日的数据为测试数据，预测出 2021 年 1 月 1 日至 2021 年 1 月 7 日的电力负荷情况，最后进行结果比较和误差分析。

本次实验硬件平台为装有 Intel i5-1035G1 处理器的台式电脑，内存为 8 GB，固态硬盘容量为 477 GB，CPU 显卡为 MX230。基于 Python 语言实现本文所提出的方法，Attention-Bi-LSTM 模型使用 Keras 深度学习框架，XGBoost 使用 py-xgboost 框架。

4.2 评价标准

本文选取平均相对误差绝对值 (mean absolute percent error, MAPE) 作为各个预测模型的主要评判

指标，并选取平均绝对误差 (mean absolute error, MAE) 和均方根误差 (root mean squared error, RMSE) 作为辅助评判指标。MAPE, MAE 和 RMSE 的计算公式分别如下所示。

$$X_{MAPE} = \frac{1}{n} \sum_{t=1}^n \frac{|y_t - \hat{y}_t|}{y_t}$$

(24)

$$X_{MAE} = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t|$$

(25)

$$X_{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2}$$

(26)

式中： n 为电力负荷数据的数量； y_t 为实际的电力负荷数据； $\hat{y}_t$ 为预测出的电力负荷数据。

4.3 数据预处理对预测结果的影响

首先对新加坡电力负荷数据使用加权灰色关联投影算法进行处理，选取与五一劳动节关联度较大的历史数据，使数据更加具有一般性。为了验证 WGRP 算法的重要性，分别使用 Attention-Bi-LSTM 和经过加权灰色关联投影算法处理的 Attention-Bi-LSTM，以及 XGBoost 和经过加权灰色关联投影算法处理的 XGBoost，对五一假期的数据进行预测。最后将上述 4 种模型的结果与真实值进行比较，并进行误差分析。

图 4 比较了引入 WGRP 算法前后 Attention-Bi-LSTM 预测结果与真实值变化趋势，可以发现经过处理的模型比未经处理的模型更接近实际值。表 1 显示了 WGRP 算法处理的 Attention-Bi-LSTM 的 MAPE, MAE 和 RMSE 值分别为 0.639, 67.147 和 96.232，意味着经过 WGRP 处理的 Attention-Bi-LSTM 预测模型具有更高的预测精度。因此，对于 Attention-Bi-LSTM，在选择 2020 年 5 月 1 日的

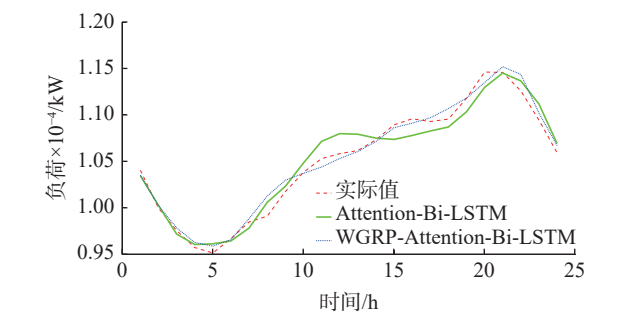


图 4 加权灰色关联投影算法预处理前后的 Attention-Bi-LSTM 预测结果与真实值对比

Fig. 4 Comparison between real values and prediction results of Attention-Bi-LSTM model before and after processing by WGRP algorithm

表 1 加权灰色关联投影算法预处理前后的 Attention-Bi-LSTM 预测误差分析

Tab.1 Error analysis of Attention-Bi-LSTM model before and after processing by WGRP algorithm

模型	$X_{MAPE} / \%$	X_{MAE} / kW	X_{RMSE} / kW
Attention-Bi-LSTM	0.971	103.063	119.726
加权灰色关联投影后的 Attention-Bi-LSTM	0.639	67.147	96.232

历史假日数据后，可以使用 WGRP 算法来提高模型的精度，表明了 WGRP 算法预处理的有效性。

图 5 比较了引入 WGRP 算法前后的 XGBoost 预测结果与真实值关系，WGRP 算法处理的 XGBoost 线条趋势与实际值基本一致，从表 2 也可以看出，WGRP 算法处理的 XGBoost 的 MAPE, MAE 和 RMSE 值分别为 0.409, 42.976 和 43.817，意味着经过 WGRP 处理的 XGBoost 模型具有更高的预测精度。因此，对于 XGBoost，WGRP 算法同样有助于提高模型的预测精度。

当不考虑 WGRP 算法时，虽然预测结果与实际值的变化趋势基本一致，但是差异很大。使用 WGRP 算法后，减少了预测结果与实际值之间的

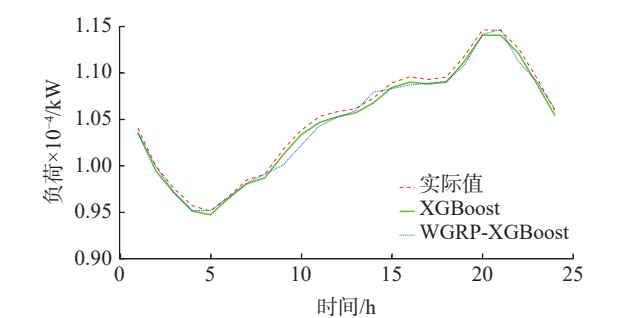


图 5 加权灰色关联投影算法预处理前后的 XGBoost 预测结果与真实值对比

Fig. 5 Comparison between real values and prediction results of XGBoost model before and after processing by WGRP algorithm

表 2 XGBoost 及加权灰色关联投影算法处理的 XGBoost 误差分析

Tab.2 Error analysis of XGBoost model before and after processing by WGRP algorithm

模型	$X_{MAPE} / \%$	X_{MAE} / kW	X_{RMSE} / kW
XGBoost	0.475	50.053	50.932
加权灰色关联投影后的 XGBoost	0.409	42.976	43.817

差距。因此, 为了减少预测误差, 首先使用 WGRP 算法处理历史数据中的假日数据。然后, 为了验证该算法在提高预测精度和减少误差方面的优势, 使用相同的模型对处理后和未处理的数据进行预测。结果发现, 数据处理的预测结果更接近真实值, 预测精度可以进一步提高。因此, 采用 WGRP 算法对假日数据进行处理, 使数据更加通用, 提高了预测精度。

4.4 预测结果分析

接下来以预处理后的电力负荷数据来验证提出的电力负荷预测方法的适用性。首先, 使用 Attention-Bi-LSTM 和 XGBoost 模型对同一组数据进行预测, 并得出两个模型的预测误差; 其次, 根据误差对上述两种预测模型进行加权, 并将其与误差倒数法相结合, 误差较小的模型被赋予更高的权重; 第三, 为了验证本文提出的组合预测模型的有效性, 使用 LSTM, Bi-LSTM, Attention-RNN, Attention-LSTM, Attention-Bi-LSTM, XGBoost 和 “Attention-Bi-LSTM + XGBoost” 组合预测模型分别进行预测, 并对预测结果进行比较。

4.4.1 预测结果和实际数据比较

a. 图 6 显示 LSTM, Bi-LSTM, Attention-RNN, Attention-LSTM, Attention-Bi-LSTM, XGBoost 模型和 “Attention-Bi-LSTM + XGBoost” 组合预测模型的预测结果与实际值之间的比较, 易知 “Attention-Bi-LSTM + XGBoost” 组合预测模型的预测精度最高, 而 Bi-LSTM 模型的预测精度最低。

b. 新加坡 2020 年 1 月 7 日 24 h 预测结果的局部放大图如图 7 所示, 从图 7 中可以看出, “Attention-Bi-LSTM + XGBoost” 组合预测模型的预测值曲线与实际值曲线最接近, 即本文提出的

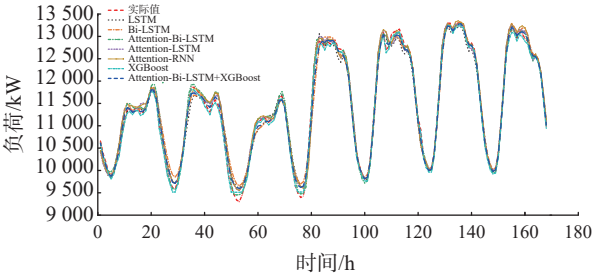


图 6 预测结果和实际数据的趋势比较图

Fig. 6 Trend comparison chart of prediction results and actual values

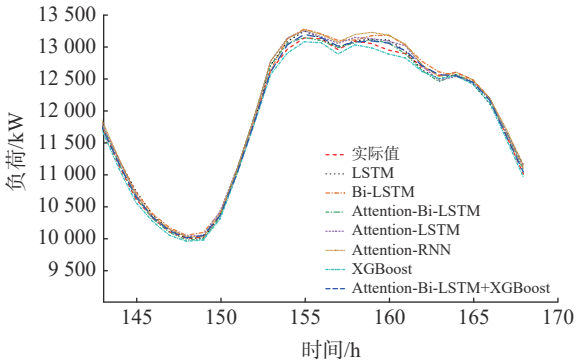


图 7 新加坡 2020 年 1 月 7 日预测结果的局部放大图

Fig. 7 Local enlarged drawing of Singapore's prediction results on January 7, 2020 (24 hours)

模型拟合效果最好。

4.4.2 误差比较分析

上述 6 种模型的 MAPE, MAE 和 RMSE 值如表 3 所示, 通过比较表 3 中误差值可知:

a. 由于 LSTM 是循环神经网络 (recurrent neural network, RNN) 的改进, 因此 Attention-LSTM 的误差小于 Attention-RNN, 说明本文选择 LSTM 作为基础预测模型的必要性。

b. 根据 LSTM, Attention-LSTM 和 Attention-

表 3 新加坡电力市场的误差分析

Tab.3 Error analysis in Singapore power market

模型	$X_{MAPE} / \%$	X_{MAE} / kW	X_{RMSE} / kW
LSTM	1.011	113.480	147.083
Bi-LSTM	1.155	127.812	168.841
Attention-RNN	0.886	98.160	130.704
Attention-LSTM	0.877	97.691	125.574
Attention-Bi-LSTM	0.711	79.418	105.023
XGBoost	0.526	59.463	62.585
combined model	0.445	49.546	63.405

RNN 模型 3 种预测模型综合比较, 可知注意力机制对预测精度有显著影响。

c. 在所有基准模型的精度测试标准中, XGBoost 的误差值最小, 这意味着 XGBoost 具有极其优异的预测性能, 因此本文使用 XGBoost 来组合提高 Attention-Bi-LSTM 预测模型的精准性。

d. 表 3 显示, “Attention-Bi-LSTM + XGBoost” 组合预测模型的 MAPE 和 MAE 值分别为 0.445 和 49.546, 从整体角度来看最小。尽管 XGBoost 的 RMSE 值最小为 62.585, 但所提出的组合预测模型的值仅略高于最小值, 显然可以接受。因此, 两种预测模型的组合可以从整体上降低预测误差, 从而优于单一预测模型, 提高现有模型的预测精度。

e. 与文献 [21] 不同的是, 本文首先对数据进行了处理, 并且将双向和注意力机制添加到 LSTM 模型进行改进, 以提高预测精度。通过误差对比, 可以发现本文所用方法相较文献 [21] 在精确度上可以提高 12% 左右, 因此证明本文使用各种改进方法的必要性和有效性。

5 结 论

为了提高电力负荷预测的准确性和稳定性, 本文提出了一种基于 WGRP 算法的 Attention-Bi-LSTM + XGBoost 电力负荷组合预测方法。在数据预处理阶段, 采用 WGRP 算法选择节假日历史负荷序列, 在预测阶段则采用 Attention-Bi-LSTM + XGBoost 电力负荷组合预测模型进行预测并得到了较好的预测结果。使用新加坡电力市场的数据集对比评估后, 可以得出以下结论:

- a. 利用 WGRP 算法对假日数据进行预处理, 可以有效提高模型的预测精度;
- b. 注意力机制允许 Bi-LSTM 模型强调重要因素的影响, 从而消除冗余, 提高预测性能;
- c. 在 XGBoost 中加入正则项可以降低模型复杂性, 有效防止过度拟合, 减少计算量, 从而大大提高算法的效率, 因此使用 XGBoost 模型进行组合优化可以极大减少模型误差;
- d. 与所有基准模型预测结果相比, “Attention-Bi-LSTM + XGBoost” 组合模型的预测结果误差最小, 与实际值最接近, 表明了本文预测方法的优越性。

综上所述, 本文提出的组合预测方法比单一

的传统和现代预测方法、混合预测方法, 以及其他现有的组合预测方法更有效, 获得更高的电力负荷预测精度, 从而减少电力市场中不必要的浪费, 提高电力系统运行的可靠性和安全性。

参考文献:

[1] 陈松伟, 任洪波, 余斌. 新电改背景下输配电价测算及传导参数分析 [J]. 上海理工大学学报, 2019, 41(2): 183-189.

[2] 潘婷婷, 李军祥, 黄芳旧. 基于需求响应的含光伏电力系统优化及其算法改进 [J]. 上海理工大学学报, 2019, 41(5): 448-454.

[3] JIANG P, LI R R, LIU N N, et al. A novel composite electricity demand forecasting framework by data processing and optimized support vector machine[J]. *Applied Energy*, 2020, 260: 114243.

[4] 葛磊蛟, 赵康, 孙永辉, 等. 基于孪生网络和长短时记忆网络结合的配电网短期负荷预测 [J]. 电力系统自动化, 2021, 45(23): 41-50.

[5] QIU X H, ZHANG L, SUGANTHAN P N, et al. Oblique random forest ensemble via least square estimation for time series forecasting[J]. *Information Sciences*, 2017, 420: 249-262.

[6] WU F, CATTANI C, SONG W Q, et al. Fractional ARIMA with an improved cuckoo search optimization for the efficient short-term power load forecasting[J]. *Alexandria Engineering Journal*, 2020, 59(5): 3111-3118.

[7] 孙庆凯, 王小君, 张义志, 等. 基于 LSTM 和多任务学习的综合能源系统多元负荷预测 [J]. 电力系统自动化, 2021, 45(5): 63-70.

[8] WANG Z Y, ZHOU X J, TIAN J T, et al. Hierarchical parameter optimization based support vector regression for power load forecasting[J]. *Sustainable Cities and Society*, 2021, 71: 102937.

[9] LIANG Y, NIU D X, HONG W C. Short term load forecasting based on feature extraction and improved general regression neural network model[J]. *Energy*, 2019, 166: 653-663.

[10] BRÉGÈRE M, HUARD M. Online hierarchical forecasting for power consumption data[J]. *International Journal of Forecasting*, 2022, 38(1): 339-351.

[11] JNR E O N, ZIGGAH Y Y, RELVAS S. Hybrid ensemble intelligent model based on wavelet transform, swarm intelligence and artificial neural network for electricity demand forecasting[J]. *Sustainable Cities and Society*, 2021, 66: 102679.

- [12] POLAMURI S R, SRINIVAS K, MOHAN A K. Multi-model generative adversarial network hybrid prediction algorithm (MMGAN-HPA) for stock market prices prediction[J]. *Journal of King Saud University-Computer and Information Sciences*, 2021, In Press.
- [13] WANG Y L, WANG L P, YANG F J, et al. Advantages of direct input-to-output connections in neural networks: the Elman network for stock index forecasting[J]. *Information Sciences*, 2021, 547: 1066–1079.
- [14] BARMAN M, CHOUDHURY N B D. A similarity based hybrid GWO-SVM method of power system load forecasting for regional special event days in anomalous load situations in Assam, India[J]. *Sustainable Cities and Society*, 2020, 61: 102311.
- [15] 刘友波, 吴浩, 刘挺坚, 等. 集成经验模态分解与深度学习的用户侧净负荷预测算法 [J]. *电力系统自动化*, 2021, 45(24): 57–64.
- [16] 赵佩, 代业明. 基于实时电价和加权灰色关联投影的 SVM 电力负荷预测 [J]. *电网技术*, 2020, 44(4): 1325–1332.
- [17] XIAO L Y, WANG J Z, HOU R, et al. A combined model based on data pre-analysis and weight coefficients optimization for electrical load forecasting[J]. *Energy*, 2015, 82: 524–549.
- [18] 林锦顺, 姚俭. 基于 BP 神经网络的组合预测及在电力负荷的应用 [J]. *上海理工大学学报*, 2005, 27(5): 451–455.
- [19] 邓带雨, 李坚, 张真源, 等. 基于 EEMD-GRU-MLR 的短期电力负荷预测 [J]. *电网技术*, 2020, 44(2): 593–602.
- [20] CHU Y Y, XU P, LI M X, et al. Short-term metropolitan-scale electric load forecasting based on load decomposition and ensemble algorithms[J]. *Energy and Buildings*, 2020, 225: 110343.
- [21] 陈振宇, 刘金波, 李晨, 等. 基于 LSTM 与 XGBoost 组合模型的超短期电力负荷预测 [J]. *电网技术*, 2020, 44(2): 614–620.
- [22] 庄家懿, 杨国华, 郑豪丰, 等. 基于多模型融合的 CNN-LSTM-XGBoost 短期电力负荷预测方法 [J]. *中国电力*, 2021, 54(5): 46–55.
- [23] NIE Y, JIANG P, ZHANG H P. A novel hybrid model based on combined preprocessing method and advanced optimization algorithm for power load forecasting[J]. *Applied Soft Computing*, 2020, 97: 106809.
- [24] ZHENG X, CHEN W Z. An attention-based bi-LSTM method for visual object classification via EEG[J]. *Biomedical Signal Processing and Control*, 2021, 63: 102174.
- [25] XIONG Y R, YANG W Y, LIAO H Y, et al. Soft variable selection combining partial least squares and attention mechanism for multivariable calibration[J]. *Chemometrics and Intelligent Laboratory Systems*, 2022, 223: 104532.

(编辑: 丁红艺)