

基于 CEEMDAN-SBiGRU-OMHA 的短期电力负荷预测^①



包广斌, 刘晨, 张波, 沈治名, 罗瞳

(兰州理工大学 计算机与通信学院, 兰州 730050)

通信作者: 刘晨, E-mail: 928871515@qq.com

摘要: 为了提高短期电力负荷预测的精准度, 充分挖掘电力负荷数据的复杂相关性, 提出了一种优化多头注意力机制的 CEEMDAN-SBiGRU 组合预测模型, 改进了特征提取和特征融合两个模块. 首先, 采用自适应噪声完全集成经验模态分解 (complete ensemble empirical mode decomposition with adaptive noise, CEEMDAN) 将电力负荷数据分解成多个内在模态函数 (IMF) 和一个残差信号 (RES); 同时引入降噪自编码器 DAE 挖掘数据中受气象因素、工作日类型和温度变化的潜在特征. 其次, 将提取到的复杂特征输入至堆叠双向门控循环单元 (stacked bidirectional gated recurrent unit, SBiGRU) 模块中继续学习, 以获取隐藏状态. 最后, 将获取的隐藏状态输入至加入残差机制和层归一化优化的多头注意力 (optimized multi-head attention, OMHA) 机制模块, 可以准确地给重要特征分配更高的权重, 解决噪声干扰问题. 实验结果表明, CEEMDAN-SBiGRU-OMHA 组合模型具有更高的精确性.

关键词: 短期电力负荷预测; 自适应噪声完全集成经验模态分解 (CEEMDAN); 堆叠双向门控循环单元 (SBiGRU); 降噪自编码器; 优化的多头注意力 (OMHA)

引用格式: 包广斌, 刘晨, 张波, 沈治名, 罗瞳. 基于 CEEMDAN-SBiGRU-OMHA 的短期电力负荷预测. 计算机系统应用, 2024, 33(10): 124-132. <http://www.c-s-a.org.cn/1003-3254/9661.html>

Short Term Power Load Forecasting Based on CEEMDAN-SBiGRU-OMHA

BAO Guang-Bin, LIU Chen, ZHANG Bo, SHEN Zhi-Ming, LUO Tong

(School of Computer and Communication, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: This study proposes a CEEMDAN-SBiGRU combined prediction model with an optimized multi-head attention mechanism to enhance the precision of short-term power load forecasting and fully explore the complex correlation of power load data. The model improves two modules: feature extraction and feature fusion. Firstly, the complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) is utilized to decompose the power load data into multiple intrinsic mode function (IMF) and a residual signal (RES); and a denoising autoencoder DAE is introduced to extract potential features from the data affected by meteorological factors, workday types, and temperature changes. Secondly, the extracted intricate features are fed into the stacked bidirectional gated recurrent unit (SBiGRU) module to obtain hidden states. Finally, the obtained hidden states are input into the optimized multi-head attention (OMHA) mechanism module, which incorporates residual mechanism and layer normalization, to accurately assign higher weights to important features and solve the problem of noise interference. The experimental results indicate that the CEEMDAN-SBiGRU-OMHA combined model achieves higher accuracy.

Key words: short term power load forecasting; complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN); stacked bidirectional gated recurrent unit (SBiGRU); noise reduction autoencoder; optimized multi-head attention (OHMA)

① 基金项目: 甘肃省自然科学基金 (18JR3RA156); 兰州市科技计划 (2017-4-105)

收稿时间: 2024-03-14; 修改时间: 2024-05-06, 2024-05-14; 采用时间: 2024-05-21; csa 在线出版时间: 2024-09-02

CNKI 网络首发时间: 2024-09-03

近年来, 鉴于保障电力系统的可靠性、电网的安全运行以及降低运营成本等多重需求^[1], 电力系统运营商对短期负荷预测的稳定性、准确性和效率提出了更为严格的标准. 在这一背景下, 短期电力负荷预测显得尤为关键. 电力负荷预测按照时间跨度的不同, 可以分为超短期、短期、中期以及长期负荷预测. 其中, 短期负荷预测专注于对接下来几小时至数天内的电力需求进行预测, 以便电力系统运营商能够有效地规划发电资源、调度能源分配和管理负荷需求.

随着计算机技术的广泛应用, 智能电网迅速发展, 在电网规划和运行中, 准确的负荷预测变得至关重要. 因此, 理解数据的内在特征和提高预测精度成为一项重要且有意义的任务. 多年来, 国内外学者们提出了不同的方法进行预测, 它们大致可以分为3类: (1) 传统预测方法; (2) 机器学习方法; (3) 深度学习方法. 传统的预测方法有回归分析^[2]、ARIMA^[3,4]、季节指数平滑^[5]、卡尔曼滤波^[6]等, 上面提出的传统方法都是基于统计学实现的, 在简单的预测模型、数据样本和高计算效率方面具有一定的优势^[7,8]. 然而, 这些方法对非线性、复杂的数据样本并不有效. 随着人工智能网络模型在许多领域取得的重大进展, 出现了机器学习算法, 如循环神经网络 (recurrent neural network, RNN)^[9]、长短期记忆 (long short term memory, LSTM)^[10]和门控循环单元 (gated recurrent unit, GRU)^[11]等, 这些技术已经在短期负荷预测中取得了良好的效果. 然而, 当单一预测模型遇到长序列、多维输入的电力负荷预测问题时, 会出现序列记忆丢失、数据结构混乱和无法深入挖掘序列潜在特征等问题, 从而影响负荷预测的准确性. 朱江行等人^[12]将 Prophet 和 XGBoost 结合的一种对电力负荷预测的模型. Khan 等人^[13]和孟磊等人^[14]使用 CNN 提取特征向量, 把处理之后的特征向量一同送入 LSTM 中进行预测. Li 等人^[15]将 CNN 提取构造的特征向量送入门控递归单元 (GRU) 进行预测, 两种模型都获得了高精度的预测值. Javed 等人^[16]提出了一种基于特征提取的 CNN-BiLSTM 组合预测模型, 充分利用电力负荷数据的时间特性, 解决了梯度消失问题, 加快了训练速度, 有效地分析和处理了电力负荷数据的潜在信息. Mounir 等人^[17]使用 EMD 算法分解原始电力负荷数据, 为了预测每个本征模态函数, 采用了混合模型 BiLSTM, 最终叠加各子序列的预测结果来进行电力负荷预测. Sideratos 等人^[18]采用 EEMD 分解的方法对负

荷数据进行分解, 对每一个时间序列的分量建立起各种 Elman 递归神经网络模型, 然而在实际的预测过程中, 调整模型参数的过程比较复杂. 因此, 为了提高短期电力负荷预测的精度, 本文提出了一种自适应噪声完全集成经验模态分解 (complete ensemble empirical mode decomposition with adaptive noise, CEEMDAN)、堆叠双向门控循环单元 (stacked bidirectional gated recurrent unit, SBiGRU) 和优化的多头注意力机制 (optimized multi-head attention, OMHA) 的组合模型. 该预测模型是在 EEMD 分解的过程中, 通过加入自适应白噪声得到 CEEMDAN 算法, 使其作为误差序列的分解算法, 获得内在模态函数 (IMF) 分量和唯一的残差信号, 它可以在少量的实验中重构错误序列, 从而提高了序列分解的效率. 使用降噪自编码器 DAE (denoising autoencoder) 有效地学习潜在特征序列, 将获取的影响特征因子与 IMF 分量合成一个新的矩阵, 从而整合了 CEEMDAN 和 DAE 的信息. 新的矩阵将作为 SBiGRU 的输入, 使 SBiGRU 同时考虑来自前向和后向序列的信息, 以学习时间序列中的复杂关系, 可以有效地处理时间序列数据, 提高了模型的综合性能. 引入多头注意力机制模块, 可以为 SBiGRU 的隐藏状态分配不同的权重, 从而增强了不同信息对负荷需求的影响. 根据实验的结果数据可以进一步得出: CEEMDAN-SBiGRU-OMHA 的组合模型在预测精度方面表现更出色.

1 模型概述

1.1 完备集合经验模态分解

Huang 等人^[19]在 1998 年提出经验模态分解 (EMD). 该方法更适用于增强复杂波动信号的平滑性, 将非平滑信号根据波动划分为 IMF. 每个 IMF 都遵循以下模式.

1) 极值点个数 i 和零点个数 v 必须满足 $i=v$ 或 $i-v=|1|$;

2) 极大值的上包络线和极小值的下包络线序列均值之和为 0.

为了解决 EMD 模态混叠的发生, EEMD 主要在原始序列中多次添加不同的高斯白噪声, 进行对信号极值点分布的改变, 然而, EEMD 方法增加的白噪声会产生冗余, 会对后面工作出现一些误差. 为了解决 EMD 和 EEMD 的问题, 完备集合经验模态分解算法在每个分解阶段添加自适应白噪声, 得到 IMF 分量以及唯一的残差信号, 可以克服 EMD 中存在的模态混叠现象,

解决了计算效率低和因增加平均次数,减少重构误差而导致 EEMD 分解不完全的问题。

记 $s(n)$ 为原始电力负荷数据时间序列, $v^i(n)$ 为第 i 次实验中添加的高斯白噪声序列, 第 i 次分解负荷数据时间序列可表示为 $s^i(n) = s(n) + v^i(n)$. 假设 EMD 和 CEEMDAN 产生的第 k 个模态分量分别表示为 $E_k(\cdot)$ 和 IMF_k . CEEMDAN 算法实现的具体步骤如下。

步骤 1: 通过 EMD 分解得到第 1 个 IMF 分量, 计算 CEEMDAN 分解得到唯一的残差信号 $r_1(n)$, 如式 (1)、式 (2) 所示:

$$IMF_1(n) = \frac{1}{I} \sum_{i=1}^I IMF_1^i(n) \quad (1)$$

$$r_1(n) = s(n) - IMF_1(n) \quad (2)$$

步骤 2: 继续获取第 2 个 IMF 分量, 如式 (3) 所示:

$$IMF_2(n) = \frac{1}{I} \sum_{i=1}^I E_1(r_1(n)) + \varepsilon_1 E_1(v^i(n)) \quad (3)$$

步骤 3: 重复上述步骤, 计算第 n 个残差信号, 如式 (4) 所示:

$$r_k(n) = r_{k-1}(n) - IMF_k(n) \quad (4)$$

则 $n+1$ 个 IMF 分量如式 (5) 所示:

$$IMF_{k+1}(n) = \frac{1}{I} \sum_{i=1}^I E_1(r_k(n)) + \varepsilon_k E_k(v^i(n)) \quad (5)$$

重复上述步骤, 等余量信号呈单调的趋势之后, 就停止分解原始信号 $s(n)$, 如式 (6) 所示:

$$s(n) = R(n) + \sum_{k=1}^K IMF_k \quad (6)$$

其中, K 为最终模态分量的数目, $R(n)$ 为最终单调残差信号。

1.2 降噪自编码器 DAE

DAE 是一种基于自动编码器 (autoencoder, AE) 优化的模型, 在逼近多变量非线性和防止过拟合问题方面具有良好的性能. 自动编码器如图 1 所示。

AE 是一种无监督学习算法, 也是对传统神经网络算法的改进, 能够自动提取小样本和未标记的特征, 因此该网络可以更好地提高特征提取的运行速度. 它在结构上由 3 层定义: 输入层、隐藏层和输出层. 虽然输入层和输出层中的节点数保持恒定, 但构成隐藏层、输入层和输出层的节点数可以根据需要而改变。

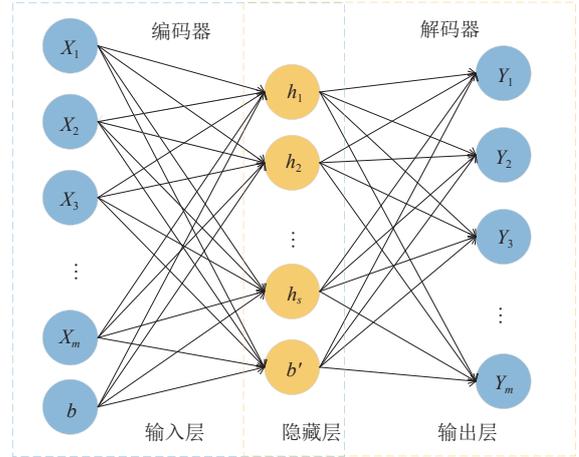


图 1 自动编码器网络结构

将原始数据作为输入 X , 经过编码器之后, 将输入 X 映射到隐藏层输出 Y , 如式 (7) 所示:

$$y = f_{\theta}(X) = S(wX + b) \quad (7)$$

其中, S 是 Sigmoid 函数, w 是编码器权值矩阵, b 是隐藏层的输入偏置向量。

经过编码得到的输出 Y , 再输入解码器后得到重构后的数据, 如式 (8) 所示:

$$Z = g_{\theta}(Y) = S(w'Y + b') \quad (8)$$

其中, w' 是解码器权值矩阵, b' 是输出层的输入偏置向量。

为了最优化提取数据的有效特征序列, 则重建的损失函数如式 (9) 所示:

$$L(X, Z) = \|X - Z\|^2 \quad (9)$$

将通过最小化平均重建误差以获得稳定的, 最优的参数, 如式 (10) 所示:

$$\min_{\theta} J_{DAE}(\theta) = \min_{\theta} \sum_{i=1}^n L(X_i, Z_i) = \min_{\theta} \sum_{i=1}^n L(X_i, g_{\theta}(f_{\theta}(X_i))) \quad (10)$$

其中, θ 为调节参数, X_i 和 Z_i 分别是第 i 次输入样本和输出样本数据, n 为输入样本的数量。

DAE 是在自编码器 AE 的基础上引入了一个加噪操作, 迫使编码器去学习输入信号, 使之更加具有鲁棒性. 这个加噪操作在输入层和隐藏层之间发挥作用, 它的引入使得编码器在训练过程中更好地捕捉了稳定的负荷变化特征因子表达, 同时确保了输出不会完全等于输入. DAE 结构图如图 2 所示。

DAE 由输入特征序列 X 、加噪特征序列 X' 、中间

特征Y和重构特征序列Z组成. DAE 的结构为编码器 f_θ 和解码器 g_θ .

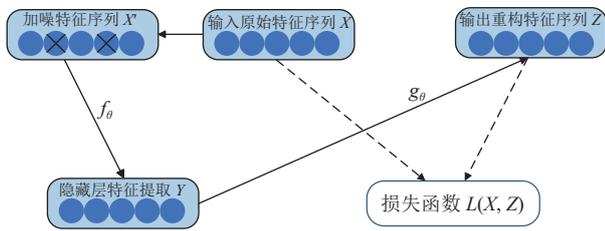


图2 DAE结构图

1.3 堆叠式双向门控循环单元

GRU 简化了 LSTM 的内部结构, 将 LSTM 里的遗忘门和输入门合并为更新门, 其主要是在前一时刻的数据上执行一定程度的记忆和有选择性的遗忘, 而重置门负责处理当前的信息并将其传递给神经网络单元, 有助于捕获时序信息中的短期依赖关系. GRU 简化了模型, 使模型参数更少, 可以解决长时间序列有效连接问题, 解决梯度爆炸或者梯度消失的问题, 对短期电力负荷预测具有一定的实用性. GRU 如图 3 所示.

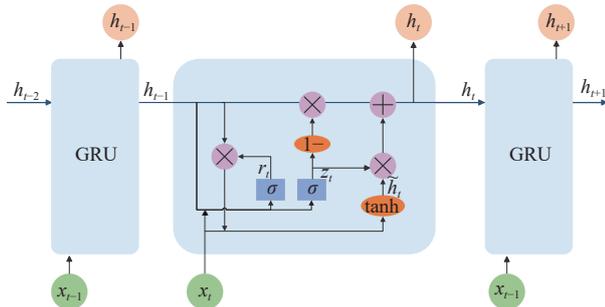


图3 GRU模型结构图

具体计算过程如式(11)–式(14)所示:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (11)$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (12)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t \times h_{t-1}, x_t]) \quad (13)$$

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \tilde{h}_t \quad (14)$$

其中, x_t 是当前输入信息, h_{t-1} 是最后保存的隐藏状态, W 是权重, σ 是 Sigmoid 激活函数, 将值压缩到 0–1 之间. 通过 \tanh 激活函数将输出值缩小到 $[-1, 1]$, 得到隐藏状态 \tilde{h}_t . r_t 用于调整输入信息 x_t 的比例. r_t 的取值范围为 0 到 1, 该值越小, 保留的输入信息越多. W_r 和 W_z 是可训练参数矩阵. z_t 用于调整历史信息保存的程度. z_t 取值从 0–1, 值越小, 保留的历史信息就越多. h_t 用于

保留需要向后传递的信息.

GRU 只能获取历史时间信息而无法获取未来时间的信息特征. 因为双向门控循环单元是由前向 GRU 和反向 GRU 组成的, 它能将隐藏层分为向前和向后两个相对的部分, 所以其可以分别读取过去和未来时间的信息. 计算公式如式(15)–式(17)所示:

$$\vec{h}_t = GRU(x_t, \vec{h}_{t+1}) \quad (15)$$

$$\overleftarrow{h}_t = GRU(x_t, \overleftarrow{h}_{t-1}) \quad (16)$$

$$h_t = W_t \vec{h}_t + V_t \overleftarrow{h}_t + b_t \quad (17)$$

其中, \vec{h}_t 表示 t 时刻的前向隐藏层状态, \overleftarrow{h}_t 表示 t 时刻的后向隐藏层状态, W_t 和 V_t 表示 \vec{h}_t 和 \overleftarrow{h}_t 对应的权重, b_t 表示 t 时刻隐藏层的偏移量. SBiGRU 由两个 BiGRU 堆叠组成, SBiGRU 的网络结构如图 4 所示.

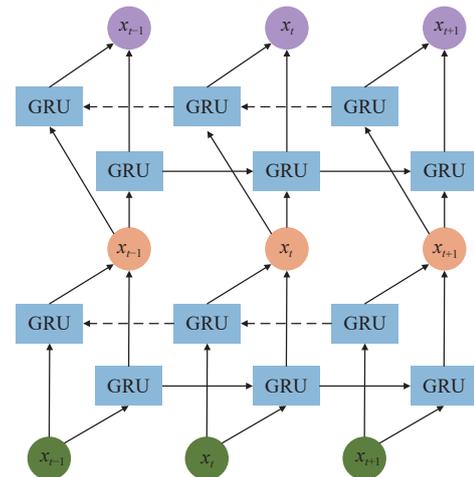


图4 SBiGRU网络结构图

1.4 优化多头注意力机制

注意力机制是受人脑思考问题时资源分配机制的启示而提出的一种深度学习方法. 在处理数据信息时, 它使用权重大小来测量不同的特征信息, 为重要特征提供较大的权重, 为相对不重要的特征提供较小的权重. 它提高了特征学习的效率, 可以从海量数据中挖掘出更有价值的隐含信息. 然而, 普通注意力机制只能从一个维度提取传感器数据特征依赖关系, 只能学习具有局限性的特征信息. 因此, 引入多头注意力机制来解决这个问题.

多头注意力机制首先通过全连接层 (FC) 将输入映射到 b 个不同的子空间. 每个子空间包含一个查询矩阵 Q , 关键矩阵 K 和价值矩阵 V . 然后, 使用缩放的点积注意力函数在 b 个子空间中并行执行注意力计算. 计算

公式如式 (18) 所示:

$$Attention(Q, K, V) = Softmax\left(\frac{Q \cdot K^T}{\sqrt{d_k}}\right) \cdot V \quad (18)$$

其中, Q 、 K 、 V 分别表示查询、键、值, d_k 表示 k 对应的维数, 计算 Q 与 K 的每个值的相似性以获得权重后, 必须将得到的权重除以 d_k , 以防止 Q 和 K 的内积过大; $Softmax(\cdot)$ 表示对得分归一化.

在多个末端并行生成多头注意力, 并将注意力的结果组合起来以获得最终输出, 计算公式如式 (19)、式 (20) 所示:

$$Multihead(Q, K, V) = Concat(head_1, \dots, head_h) W^O \quad (19)$$

$$head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \quad (20)$$

其中, $i=1, 2, \dots, h$; W_i^Q, W_i^K, W_i^V 为预测矩阵, $Multihead(\cdot)$ 表示 h 个头的注意力输出相加.

为了解决 SBiGRU 学习序列数据时, 会出现效率低、时间长和局部信息丢失等问题, 本文引入了残差机制和层归一化优化的多头注意力机制, 模型结构的简化提高了训练速度, 可以帮助模型有效地学习强相关因素的影响, 还去除了掩码矩阵操作和前馈层的计算过程. 优化后的多头注意力机制提高了预测精度, 计算公式如式 (21)、式 (22) 所示:

$$y_t = Multihead(Q, K, V) + h_t \quad (21)$$

$$E = Normalization(y_t) \quad (22)$$

其中, h_t 为多头注意力机制的输入; y_t 是经过优化后的多头注意力机制的输出; E 为 y_t 归一化之后的输出.

优化后的多头注意力机制的结构如图 5 所示.

1.5 CEEMDAN-SBiGRU-OMHA 模型总框架

本文提出一种 CEEMDAN-SBiGRU-OMHA 短期电力负荷预测模型. 模型总框架如图 6 所示.

1) 使用 CEEMDAN 算法将负荷数据分解为多个 IMF 分量和唯一残差 RES.

2) 由于 IMF 数量较多, 引入降噪自编码器 DAE 让其有效地学习更多的潜在特征序列, 可以帮助模型识别和提取潜在的、对负荷需求具有影响的特征, 从而优化数据的质量.

3) 将提取相应的特征矩阵输入到 SBiGRU 层进行训练, 这种两层 BiGRU 不仅可以为每个时间点的负荷预测提供过去和未来的负荷信息, 而且通过其多层网络设计结构提高了模型的泛化能力.

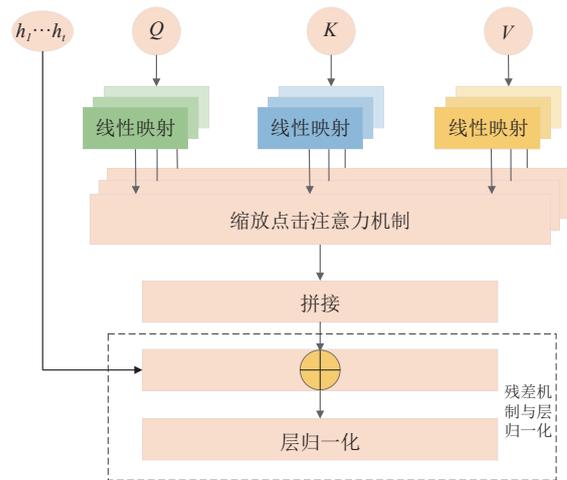


图 5 优化后的多头注意力机制的结构图

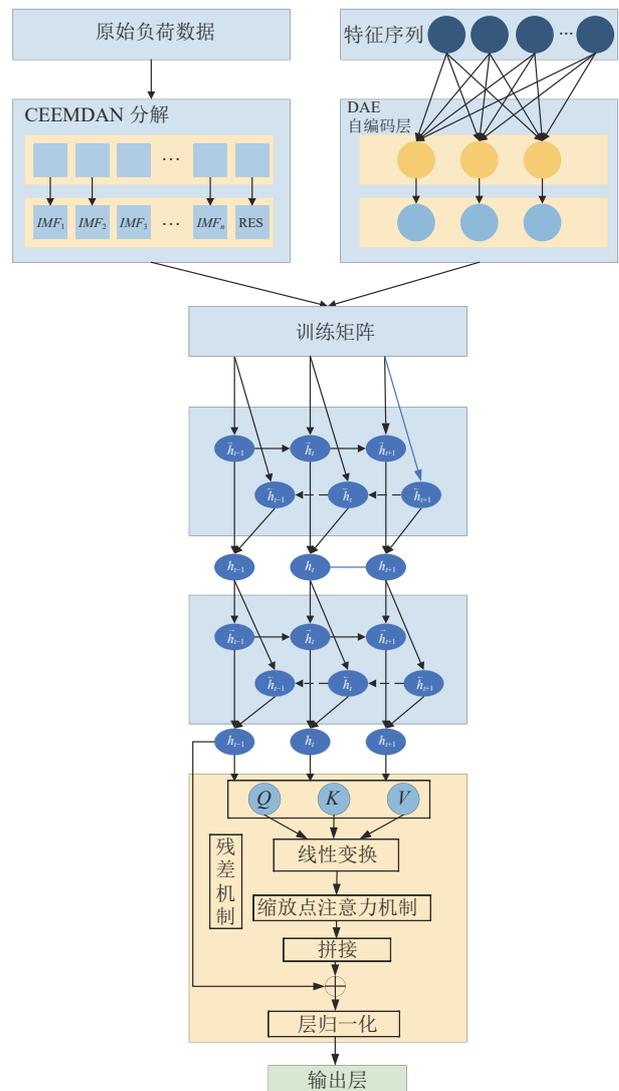


图 6 CEEMDAN-SBiGRU-OMHA 模型总框架

4) SBiGRU 网络层提取的特征用于优化多头注意力机制层的查询 (Q)、键 (K)、值 (V)。优化后的多头注意力机制层对这些查询、键、值进行线性变换、缩放点积注意力机制和拼接处理。根据权重的分配原理计算不同特征向量对应的概率, 随后迭代更新权重参数矩阵以提升模型的效果。为了应对数据的残缺问题, 将 SBiGRU 层的输出与经优化后的多头注意力机制层的输出相加并进行层归一化。

5) 将优化多头注意力机制训练好的矩阵输入到输出层, 通过全连接层输出最终的电力负荷预测值。

2 算例分析

2.1 实验数据集

本文的实验研究是选取澳大利亚某地区的电力负荷数据, 数据集包含历史负荷、日平均气温和节假日数据, 负荷数据采样周期为 3 h, 样本数为 19 808 个, 其中前 19 500 个负荷数据作为验证集, 后 308 个负荷数据作为测试集。

2.2 数据预处理

在实验之前, 我们使用归一化对样本数据进行预处理, 将其限制在 $[0, 1]$ 的范围内, 以消除奇异样本数据对预测精度的负面影响。归一化如式 (23) 所示:

$$x_c = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (23)$$

其中, x_c 为归一化的负荷数据, x_i 为原始负荷数据, x_{\max} 和 x_{\min} 分别为数据的最大值和最小值。

2.3 评价指标

本研究采用均方根误差 ($RMSE$)、平均绝对百分比误差 ($MAPE$) 和相关系数 (R^2) 作为预测的误差评价指标。评价指标计算如式 (24)–式 (26) 所示:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \hat{x}_i)^2} \quad (24)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{x_i - \hat{x}_i}{x_i} \right| \times 100\% \quad (25)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (26)$$

其中, x_i 为第 i 个预测样本的实际值, \hat{x}_i 为预测值, \bar{x} 为预测样本均值, n 为样本总数。 $RMSE$ 和 $MAPE$ 值越小, 预测效果越好, R^2 值越趋向于 1, 拟合程度越好。

3 实验研究

3.1 实验设置

实验代码基于 PyTorch 深度学习框架编写, 并在 PyCharm 代码编辑器上运行。训练过程使用显存为 24 GB 的 GPU (NVIDIA GeForce RTX 3090)。具体实验参数设置如表 1 所示。

表 1 实验中相关参数设置

参数	取值
epoch	100
学习率	0.001
激活函数	ReLU
优化器	Adam
输入序列步长	128
输出序列步长	1

3.2 CEEMDAN 的分解

当余量信号 $R(n)$ 变为常数或单调函数时, 经验模态分解 (EMD) 过程终止, 从而确定最终的本征模态函数 (IMF) 分量个数。在对负荷数据进行分解后, 得到 11 个 IMF 分量和 1 个 RES 趋势项。结果如图 7 所示, 各个分量的频率从高到低依次分布, 低频分量的波动更加规律和平缓。每个 IMF 分量代表了不同时间尺度上的特征, 使模型能够充分挖掘数据特征, 同时避免了各个局部特征之间的相互干扰, 最终的残差分量反映了负荷的总体趋势。

3.3 实验结果分析

3.3.1 对比实验

为了验证本文模型的有效性, 将其与以下 5 个模型进行比较, 结果如表 2 所示。

相比于以上 5 种深度学习模型, 本文所提出的 CEEMDAN-SBiGRU-OMHA 组合模型在短期电力负荷预测方面具有优势。从表 2 中可以看出, 单一模型的预测精度无明显区别, 本文提出的模型相比其他模型: $RMSE$ 值分别优化了 41.72%、39.66%、38.24%、30.94%、29.96%; $MAPE$ 值分别优化了 73.89%、71.49%、61.28%、38.96%、31.33%; R^2 值分别优化了 4.20%、2.60%、1.96%、1.23%、0.51%。CEEMDAN-SBiGRU-OMHA 能够提高预测结果的主要原因在于:

SBiGRU 的高效序列处理能力更全面地捕捉序列中的依赖关系. 同时, 引入优化后的多头注意力机制有效提

高了对关键信息的捕捉能力, 从而在电力负荷预测任务中表现更为卓越.

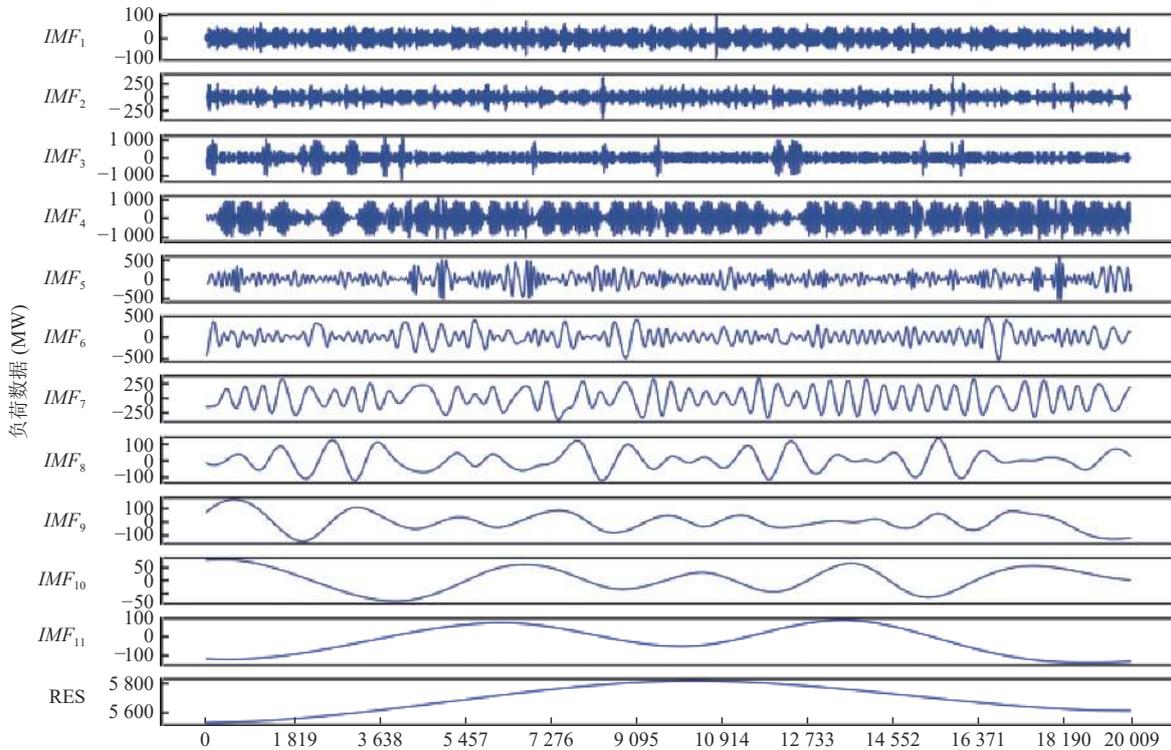


图7 CEEMDAN 分解结果图

表2 对比实验的评价指标

模型	RMSE	MAPE (%)	R ²
LSTM	18.026	7.915	0.951
GRU	17.409	7.251	0.966
LSTM-AE-CNN	17.011	5.338	0.972
SRDCC-BiLSTM	15.212	3.386	0.979
EMD-Bi-LSTM	14.997	3.010	0.986
CEEMDAN-SBiGRU-OMHA	10.505	2.067	0.991

3.3.2 消融实验

为了进一步说明本文提出模型的有效性, 还与 CEEMDAN-SBiGRU 和 EEMD-SBiGRU-Attention 模型进行了比较, 结果如图 8-图 10 所示.

图 8-图 10 表明, 在 24h 时间间隔内, 基于 CEEMDAN-SBiGRU-OMHA 组合模型在 3 种评价指标上均优于 EEMD-SBiGRU-Attention, CEEMDAN-SBiGRU 模型. 其中图 8 和图 9 表明, CEEMDAN-SBiGRU-OMHA 模型在 RMSE 与 MAPE 上相比 EEMD-SBiGRU-Attention 模型分别平均降低了 1.5% 与 0.2%, 相比 CEEMDAN-SBiGRU 模型分别平均降低了 4% 与 0.7%, 可见, 通过添加优化后的多头注意力机制, 可以更好地提高捕捉

关键信息的能力; 利用 CEEMDAN 方法对非光滑、非线性的原始数据进行分解, 可以有效地提取光滑信息; 此外, 通过在预测过程中利用这种平滑信息的叠加增强预测结果, 提高预测的准确性. 图 10 表明, CEEMDAN-SBiGRU-OMHA 预测模型的 R² 值整体优于其他两个模型的值, 分别提高了 12% 与 28%, 这说明本文提出的组合模型对于预测精度有明显的提升.

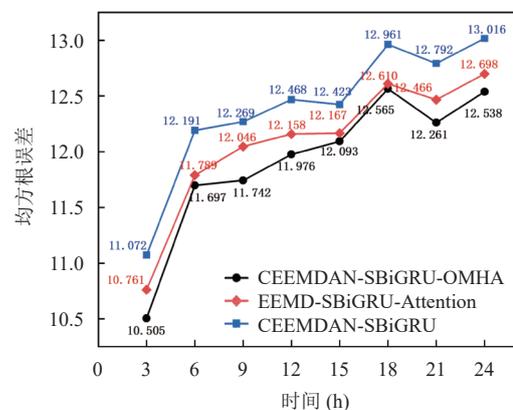


图8 3种模型均方根误差预测比较

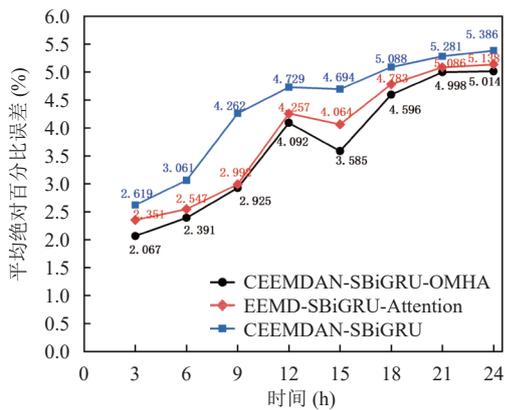


图9 3种模型平均绝对百分比误差比较

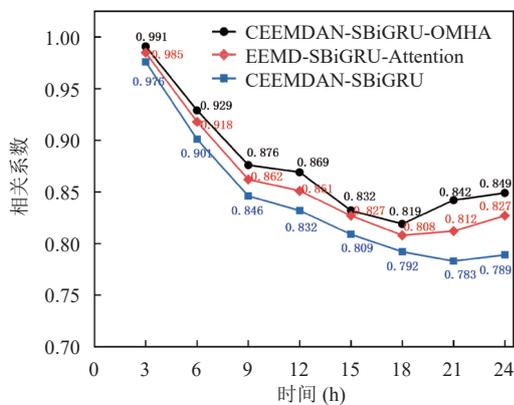


图10 3种模型相关系数比较

4 结论

实验结果表明, CEEMDAN-SBiGRU-OMHA 在预测短期电力负荷方面表现出色, 其预测误差最小, 预测精度最高. 并与 LSTM、GRU、LSTM-AE-CNN、SRDCC-BiLSTM 和 EMD-BI-LSTM 模型进行了比较, 基于实验结果, 可以得出以下结论: 1) CEEMDAN-SBiGRU-OMHA 的 3 个评估指标均优于其他预测模型, 表明本文提出的预测模型具有更好的性能; 2) 通过使用 CEEMDAN 方法, 将误差序列分解为多个更稳定的分量, 以更精确地捕捉电力负荷的复杂性和波动性; 3) 利用降噪自编码器 DAE 能够学习更多的潜在特征序列, 将这些特征因子加入分解得到的 *IMF* 分量中. 这有助于更全面地考虑不同的因素对电力负荷的影响, 有效地去除数据中的干扰噪声, 从而提高预测的综合性能; 4) 通过引入 SBiGRU 模型, 为每个分量建立 SBiGRU 预测模型, 以提高更稳定分量的误差预测精度. 这有助于减少误差并提高模型的稳定性; 5) 引入了

残差机制和层归一化改进的多头注意力机制, 以提升对负荷数据的处理效率. 这一改进有效解决了重要信息丢失的问题, 保障了关键特征的完整传递, 从而提高了预测准确度.

参考文献

- 1 王凌云, 林跃涵, 童华敏, 等. 基于改进 Apriori 关联分析及 MFOLSTM 算法的短期负荷预测. 电力系统保护与控制, 2021, 49(20): 74–81.
- 2 Yildiz B, Bilbao JI, Sproul AB. A review and analysis of regression and machine learning models on commercial building electricity load forecasting. Renewable and Sustainable Energy Reviews, 2017, 73: 1104–1122. [doi: 10.1016/j.rser.2017.02.023]
- 3 Bessani M, Massignan JAD, Santos TMO, *et al.* Multiple households very short-term load forecasting using Bayesian networks. Electric Power Systems Research, 2020, 189: 106733. [doi: 10.1016/j.epr.2020.106733]
- 4 Pierre AA, Akim SA, Semenyko AK, *et al.* Peak electrical energy consumption prediction by ARIMA, LSTM, GRU, ARIMA-LSTM and ARIMA-GRU approaches. Energies, 2023, 16(12): 4739. [doi: 10.3390/en16124739]
- 5 Deng CR, Zhang XY, Huang YM, *et al.* Equipping seasonal exponential smoothing models with particle swarm optimization algorithm for electricity consumption forecasting. Energies, 2021, 14(13): 4036. [doi: 10.3390/en14134036]
- 6 Zheng TX, Girgis AA, Makram EB. A hybrid wavelet-Kalman filter method for load forecasting. Electric Power Systems Research, 2000, 54(1): 11–17. [doi: 10.1016/S0378-7796(99)00063-2]
- 7 Cai GW, Wang WJ, Lu JH. A novel hybrid short term load forecasting model considering the error of numerical weather prediction. Energies, 2016, 9(12): 994. [doi: 10.3390/en9120994]
- 8 Paparoditis E, Sapatinas T. Short-term load forecasting: The similar shape functional time-series predictor. IEEE Transactions on Power Systems, 2013, 28(4): 3818–3825. [doi: 10.1109/TPWRS.2013.2272326]
- 9 邓斌, 张楠, 王江, 等. 基于 LTC-RNN 模型的中长期电力负荷预测方法. 天津大学学报(自然科学与工程技术版), 2022, 55(10): 1026–1033.
- 10 高超, 孙谊娟, 赵洪峰, 等. 基于 ICOA-LSTM 的短期负荷预测研究. 电子测量技术, 2022, 45(13): 88–95.
- 11 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. 电网技术, 2020, 44(9): 3416–3423.

- 12 朱江行, 邹晓松, 熊炜, 等. 基于 Prophet 与 XGBoost 混合模型的短期负荷预测. 现代电力, 2021, 38(3): 325–331.
- 13 Khan ZA, Hussain T, Ullah A, *et al.* Towards efficient electricity forecasting in residential and commercial buildings: A novel hybrid CNN with a LSTM-AE based framework. Sensors, 2020, 20(5): 1399. [doi: [10.3390/s20051399](https://doi.org/10.3390/s20051399)]
- 14 孟磊. 基于 CNN-LSTM 网络的电力负荷预测. 机电信息, 2023(7): 7–10.
- 15 Li C, Li GJ, Wang KY, *et al.* A multi-energy load forecasting method based on parallel architecture CNN-GRU and transfer learning for data deficient integrated energy systems. Energy, 2022, 259: 124967. [doi: [10.1016/j.energy.2022.124967](https://doi.org/10.1016/j.energy.2022.124967)]
- 16 Javed U, Ijaz K, Jawad M, *et al.* A novel short receptive field based dilated causal convolutional network integrated with bidirectional LSTM for short-term load forecasting. Expert Systems with Applications, 2022, 205: 117689. [doi: [10.1016/j.eswa.2022.117689](https://doi.org/10.1016/j.eswa.2022.117689)]
- 17 Mounir N, Ouadi H, Jrhilifa I. Short-term electric load forecasting using an EMD-BI-LSTM approach for smart grid energy management system. Energy and Buildings, 2023, 288: 113022. [doi: [10.1016/j.enbuild.2023.113022](https://doi.org/10.1016/j.enbuild.2023.113022)]
- 18 Sideratos G, Ikonomopoulos A, Hatziargyriou ND. A novel fuzzy-based ensemble model for load forecasting using hybrid deep neural networks. Electric Power Systems Research, 2020, 178: 106025. [doi: [10.1016/j.epsr.2019.106025](https://doi.org/10.1016/j.epsr.2019.106025)]
- 19 Wu ZH, Huang NE. Ensemble empirical mode decomposition: A noise-assisted data analysis method. Advances in Adaptive Data Analysis, 2009, 1(1): 1–41. [doi: [10.1142/S1793536909000047](https://doi.org/10.1142/S1793536909000047)]

(校对责编: 张重毅)