

引用格式: 刘杰, 马子健, 周博文, 等. 基于改进灰狼优化算法的 TCN-BiGRU 电力负荷预测 [J]. 电子科技大学学报, 2025, 54(6): 916-923.
LIU J, MA Z J, ZHOU B W, et al. TCN-BiGRU power load prediction based on improved gray wolf optimization algorithm[J]. Journal of University of Electronic Science and Technology of China, 2025, 54(6): 916-923.

基于改进灰狼优化算法的 TCN-BiGRU 电力负荷预测



刘杰*, 马子健, 周博文, 吴海滨

(哈尔滨理工大学 测控技术与通信工程学院, 哈尔滨 150080)

摘要: 为了提高短期电力负荷预测精度, 该文提出了一种基于改进灰狼优化算法的 TCN-BiGRU 模型。输入序列先由改进后的时间卷积网络 (TCN) 捕捉长期依赖关系, 再通过改进自注意力优化的双向门控循环单元 (BiGRU) 提取双向依赖关系。在模型内部结合 AR 模块、选举机制提升预测准确性, 最后通过改进的灰狼优化算法优化 TCN-BiGRU 模型的参数以提升模型的综合性能。实验仿真表明, 该模型的 MAPE、MAE 和 RMSE 分别为 4.974%、0.029、0.034, 均优于主流对照模型, 有效提升了负荷预测精度。

关键词: 短期负荷预测; 电力系统; 灰狼优化算法; 自注意力机制; 组合模型

中图分类号: TP391 **文献标志码:** A **DOI:** 10.12178/1001-0548.2024335

TCN-BiGRU power load prediction based on improved gray wolf optimization algorithm

LIU Jie*, MA Zijian, ZHOU Bowen, and WU Haibin

(College of Measurement and Control Technology and Communication Engineering, Harbin University of Science and Technology, Harbin 150080, China)

Abstract: To improve the accuracy of short-term power load forecasting, this paper proposes a TCN-BiGRU model based on an improved grey wolf optimization algorithm. In this framework, the input sequence is first processed by an enhanced temporal convolutional network (TCN) to capture long-term dependencies, and then by an improved self-attention-optimized bidirectional gated recurrent unit (BiGRU) to extract bidirectional dependencies. An auto regression (AR) module and an election mechanism are integrated within the model to enhance forecasting accuracy. Finally, the model parameters of the TCN-BiGRU are optimized using the improved grey wolf optimization algorithm to further boost its overall performance. Experimental simulations demonstrate that the proposed model achieves a mean absolute percentage error (MAPE) of 4.974%, mean absolute error (MAE) of 0.029, and root mean square error (RMSE) of 0.034, outperforming mainstream benchmark models and effectively enhancing load forecasting accuracy.

Key words: short-term load forecasting; power system; wolf optimizer algorithm; self-attention mechanism; combinatorial model

电力负荷预测方法可以归纳为经典预测方法、传统预测方法和现代智能预测方法^[1]。经典预测方法通常依赖专业人员长期积累的经验, 利用各种指标如单耗法、电力弹性系数法、负荷密度法等, 对未来电力负荷的走向进行推测。传统预测方法主要有趋势外推法、回归分析法、时间序列法等手段。现代智能预测方法主要通过专家系统^[2]、支持向量

机 (support vector machine, SVM)^[3]、神经网络 (artificial neural network, ANN)^[4]、Transformer^[5]、组合模型^[6-7]等技术来进行预测。

时间卷积网络 (temporal convolutional network, TCN)^[8] 能够并行处理数据, 因此在效率上具有显著优势。此外, TCN 继承了传统卷积神经网络的参数共享特性, 有效避免了梯度消失和梯度爆炸的

收稿日期: 2024-12-11

基金项目: 黑龙江省自然科学基金 (LH2023E086); 黑龙江省交通运输厅科技项目 (HJK2024B002)

作者简介: 刘杰, 博士, 副教授, 主要从事人工智能与图像处理、大数据模型及预测方面的研究。

*通信作者 E-mail: liujie@hrbust.edu.cn

问题。通过采用因果卷积和扩张卷积, TCN 能够捕捉更长时间范围内的依赖关系, 并防止信息丢失。TCN 的提出为负荷预测提供了一种更加优越的解决方案。

组合模型将多个预测模型结合在一起, 可以综合考虑各模型的优势和不足, 提高短期负荷预测的整体性能^[9]。文献 [10] 提出了一种结合卷积神经网络 (convolutional neural networks, CNN) 和长短时记忆网络 (long short-term memory, LSTM) 的负荷预测框架, 利用 CNN 层提取输入数据特征, 利用 LSTM 层进行序列学习, 所提模型性能得到了明显提升。文献 [11] 提出了一种包含门控循环单元 (gated recurrent unit, GRU) 和 TCN 的 GRU-TCN 交互学习方法, 用于对电网数据进行预测, 实验结果显示, 提出的网络即使在信息不完整的情况下也可以捕获到耦合关系, 并准确预测多能源交互数据, 但它只从一个方向对输入序列进行编码和解码, 导致对序列中某些关键依赖关系的理解不足。

综上所述, 短期电力负荷数据具有波动性大、易受外界因素影响以及非线性的特点, 经典的负荷预测方法虽然在一些情况下具有较高的准确性, 但通常依赖于经验和固定的指标, 难以适应复杂的电力负荷变化趋势。传统预测方法在一定程度上克服了这些问题, 但依然受到线性假设和模型局限性的制约。现代智能预测方法通过结合先进的机器学习技术, 提供了更为准确的负荷预测。然而, 现有的组合模型、卷积神经网络与长短时记忆网络的结合等方法, 虽然在一

定程度上提高了预测精度, 但仍存在信息捕捉不完全、序列依赖关系的处理能力有限的问题。

因此本文在此基础上搭建了基于 TCN-BiGRU 的组合模型, 具体而言, 多尺度卷积的 TCN 能够更好地捕捉电力负荷数据中的局部信息, 避免信息丢失, 融合改进 ISA 机制的 BiGRU 能够捕捉更多针对电力负荷数据的区间信息, 结合 AR 模块与选举机制克服负荷数据的波动性, 通过引入这一混合模型, 本文能够更好地处理多变的电力负荷数据, 捕捉复杂的非线性依赖关系, 提升预测精度。

1 模型构建

1.1 TCN-BiGRU 模型

本文建立的 TCN-BiGRU 模型融合了改进的 TCN 和 BiGRU 结构, 其中 TCN 部分用于提取特征和网络隐藏信息, BiGRU 部分则负责特征识别和时序数据匹配。此外, 模型结合了选举机制和自回归模型, 对时序数据进行深度处理, 增加模型的预测准确率。

图 1 展示了 TCN-BiGRU 模型的整体框架, 主要包括输入层、改进的 TCN 层、优化的 BiGRU 结构、AR 模块和选举机制。输入序列 X 首先根据输入步长分为若干份, 以 $X^{(1)}$ 为例, 输入后先由 TCN 捕捉长期依赖关系, 得到时间序列 X_1, X_2, \dots, X_n 后再通过 ISA 优化的 BiGRU 提取双向依赖。AR 模块通过自回归特性预测数据, 保留线性与非线性关系, 有效还原原始负荷数据, 选举机制 Y 提升预测准确性, 最终输出预测值 \hat{Y}_t 。

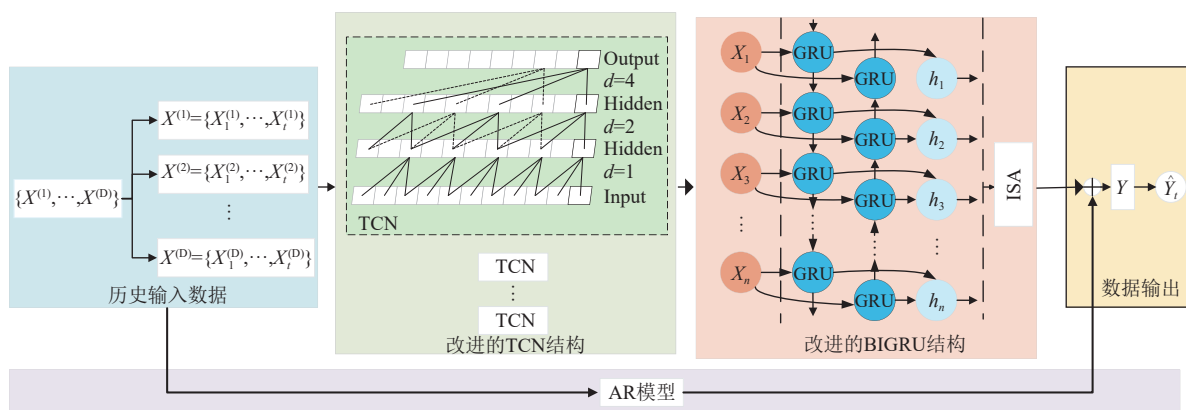


图 1 TCN-BiGRU 模型结构图

1.2 TCN 的改进

TCN 是一种专门处理序列数据的深度学习模型, 很好地解决了负荷数据中卷积核的尺寸对卷积网络提取能力限制的问题, 但其仍存在不足之处。文献 [12] 指出, TCN 在膨胀因果卷积后, 大量的

局部信息丢失, 进而限制了预测性能。故本文对 TCN 的卷积结构、残差结构和激活函数进行改进, 进一步提高短期负荷预测精度。

原有的 TCN 仅提取单一时间尺度的信息, 导致提取效率和能力受限。为此, 本文设计了多尺度

卷积结构，使用不同大小的卷积核分别捕捉长周期的全局信息和短周期的局部信息，并通过并行拼接实现特征融合，最终输出数据信息。其运算过程如下：

$$\begin{cases} S_i = C_i(x) \\ S = \text{Concat}(S_i) \end{cases} \quad (1)$$

式中， x 表示输入数据； C_i 表示卷积核大小为 i 的一维卷积运算； S_i 表示卷积之后的输出状态；Concat 表示并行拼接操作； S 表示拼接后的输出结果。

本文采用 3 种尺度的卷积核来捕获数据的分布特征，其结构图如图 2 所示。

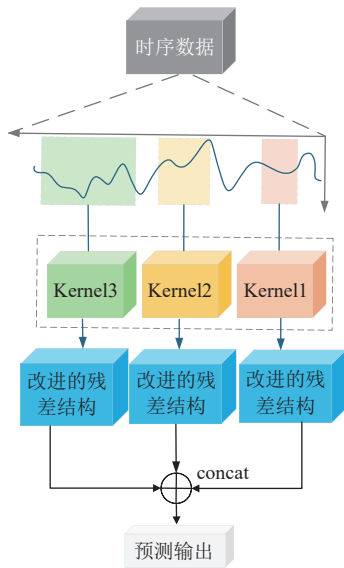


图 2 多尺度 TCN 卷积结构

由图 2 可以看出，在多尺度 TCN 卷积结构中，Kernel1、Kernel2、Kernel3 表示不同大小的卷积核，卷积完成后送入本文改进的残差结构，并通过 concat 对不同尺度特征进行拼接，将融合后的结果进行输出。

TCN 中的残差连接可以实现信息的跨层传递，但这样的连接方式会使模型的学习能力受限，因此，本文对 TCN 的残差结构进行了改进，改进后的结构如图 3 所示。将串行连接的方式改为并行拼接，又在 Dropout 后同时引入最大池化和平均池化给模型增加非对称因素，从而使网络学习到更多有用信息，增强了网络的特征提取能力，两池化的输出表示为：

$$\begin{cases} y_m = \text{GELU}(W_m \cdot x + b_m) \\ y_a = \text{GELU}(W_a \cdot x + b_a) \end{cases} \quad (2)$$

式中， y_m 和 y_a 分别为最大池化和平均池化的输出； W_m 和 W_a 分别为各支路上进行的卷积和池化操作； b_m 和 b_a 为各支路的加性偏置项。

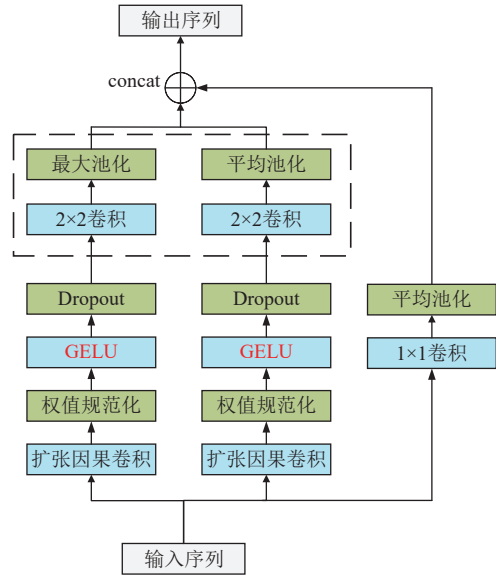


图 3 改进的残差结构

原始的 TCN 选用 ReLU 作为激活函数，该激活函数在 $x > 0$ 时直接输出数据的全部信息，反之输出为 0，导致神经元停止学习并引发梯度消失。为了改进此问题，选用 GELU 对 ReLU 进行替换。GELU 函数的表达式近似为：

$$\text{GELU}(X) = 0.5x \left(1 + \tanh \left[\sqrt{\frac{2}{\pi}} (x + 0.044715x^3) \right] \right) \quad (3)$$

式中， x 为输入数值；tanh 为双曲正切函数。

当输入 $x > 0$ 时，GELU 激活函数图像近似为直线；当 $x < 0$ 时，GELU 激活函数具有非线性输出的能力。因此选用 GELU 激活函数对时序信号进行特征提取。

1.3 ISA 机制改进 BiGRU

现有的自注意力 (self-attention, SA) 机制经常使用不同的映射矩阵将 Q (查询语句向量矩阵)、 K (关键字向量矩阵)、 V (值向量矩阵) 映射到不同的特征空间中，本文改进的 ISA 机制将 Q 和 K 映射到同一特征空间， V 映射到高维空间以提取关键信息。另外，原始 SA 在计算编码信息时总是以单点计算的方式计算最细粒度上的信息，而实际的电力负荷序列总是以多点区间集合的形式存在，因此本文在原有的基础上提出了多粒度机制来捕获更多区间的相关信息，获得更多关于短期电力负荷数据的非线性特征，公式表示如下：

$$\text{Attention}(Q, K, V) = \sum_{i=1}^n \text{Softmax} \left(\frac{Q_i K_i^T}{\sqrt{d_k}} \right) V_i \quad (4)$$

式中， Q_i 、 K_i 、 V_i 分别为每一粒度映射后的查询矩阵、键矩阵、值矩阵。

通过提出的多粒度机制, 可以捕获时间序列更广泛的信息, 同时保留细节信息, 进而提高时序任务的处理能力, 其结构示意图如图 4 所示, W_1 、 W_2 代表本文设定的两种不同粒度, 将 ISA 机制中的查询矩阵 Q 和键矩阵 K 映射到相同特征空间, 并通过 Softmax 函数归一化权重。在数据集分析中, 粒度分别设为 1 h 和 10 h, 最大限度地保留序列的低层和高层信息, 提升预测效果。

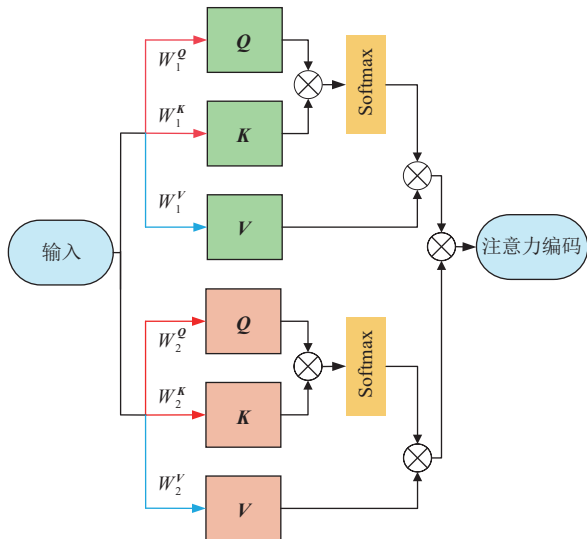


图 4 ISA 机制运算示意图

结合 ISA 机制对 BiGRU 进行改进, 使得每个 GRU 单元都能对权重信息进行处理, 改进后的模型结构图如图 5 所示, 在图 5 中, 历史数据首先通过 BiGRU 进行特征学习并输出隐藏状态。ISA 动态调整时间步权重, 对 BiGRU 的隐藏状态进行加权平均, 帮助 BiGRU 更有效地处理数据, 提升整体预测效果。

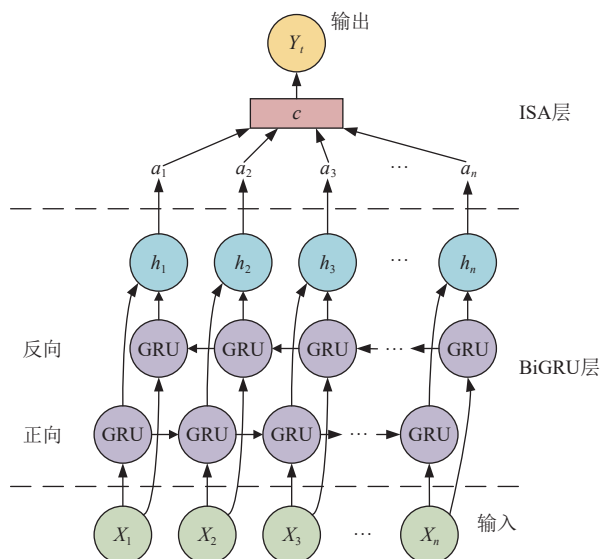


图 5 ISA 机制改进 BiGRU 结构图

1.4 AR 模块与选举机制

针对短期电力负荷数据波动性强、外部扰动因素影响较大的特点, 本文在模型中结合了 AR 模块与选举机制来提升预测准确性。

由于负荷数据具有明显的周期性, 因此选择在模型中引入 AR 模块, AR 模块通过自回归特性预测数据, 提取负荷数据中的线性关系, 能够有效还原原始负荷数据。其预测结果如下:

$$l_t = \phi \cdot x + b \quad (5)$$

式中, ϕ 和 b 分别表示自回归参数矩阵和偏置项; x 为原始输入序列; l_t 为 AR 模块预测结果。

将 AR 模块预测结果和其他非线性模型预测结果进行整合, 得到最终的预测结果, 表示为:

$$Y_t = \alpha d_t + \beta l_t \quad (6)$$

式中, d_t 为非线性模型的预测结果; α 和 β 表示两种预测结果所占的权重; Y_t 为模型在 t 时刻最终的预测结果。

本文提出的选举机制将历史数据和预测值视为选民, 赋予投票权, 根据各自权重决定预测推荐值。通过选举机制可以从数据的整体分布和个体点的相似性对数据进行分析, 提供了一种提升组合预测模型性能的方法。通过定义选举函数, 进一步体现选举机制的优势, 定义为:

$$Y = \alpha \text{His} + \beta \text{Pre} \quad (7)$$

式中, α 、 β 分别为历史值和预测值所占的比重; His、Pre 分别表示数据的历史值和预测值; Y 表示数值权重融合后模型的推荐值。

由式 (7) 可知, α 、 β 的取值对于推荐值的确定较为关键, 通过定义目标函数 $F(\alpha, \beta)$ 来实现对参数的优化, 其公式为:

$$F(\alpha, \beta) = \sqrt{\frac{1}{N} \sum_{i=1}^N (\alpha \text{His}_i + \beta \text{Pre}_i - Z_i)^2} \quad (8)$$

式中, N 为实验中的测试样本数; Z_i 表示第 i 个数据当前的真实值。

1.5 灰狼优化算法的改进

GWO 是一种模拟灰狼社会等级和捕猎机制的群智能优化算法^[13], 具有自组织性强、收敛速度快和参数少等优点, 常用于参数优化。然而, GWO 可能陷入局部最优解, 因此本文在种群质量、控制系数和位置更新等关键方面对其进行了改进。并使用改进后的 IWGO 算法确定 TCN-BiGRU 模型的卷

积核大小、学习率、迭代次数、隐藏层节点数、隐藏层数，以此最优化 TCN-BiGRU 模型。

种群质量对优化结果影响显著。因此，本文采用 Fuch 混沌映射方法初始化灰狼种群。其数学表达式如下：

$$h(v+1) = \cos \left[\frac{1}{h(v)^2} \right] \quad h(v) \neq 0, v = 1, 2, \dots, T \quad (9)$$

利用透镜成像反学习机制，在 Fuch 混沌映射的搜索空间中映射另一组空间，扩大搜索范围，选择最优空间，增强狼群多样性，提升种群质量，避免初始阶段陷入局部最优。其公式表示为：

$$x_i^* = \frac{a_i + b_i}{2} + \frac{a_i + b_i}{2k} - \frac{x_i}{k} \quad (10)$$

式中， a_i 和 b_i 为搜索空间第 i 维的最低值和最高值； x_i 为当前个体在第 i 维的分量； x_i^* 为个体的透镜反解； k 为比例因子，本文取 7 000。

在 GWO 中，距离控制系数 a 表示为：

$$a = 2 - 2 \frac{t}{t_{\max}} \quad (11)$$

式中， t 为当前迭代次数； t_{\max} 为最大迭代次数。

式 (11) 对新解的搜寻能力较弱，易陷入局部最优。故对其进行修改，帮助算法跳出局部最优，公式表示为：

$$a = 1 + \sin \left(\frac{\pi}{2} + \pi \left(\frac{t}{t_{\max}} \right) \right) \quad (12)$$

改进后的 a 引入非线性收敛因子来平衡其收敛性。 a 的波动性得到了提升，综合考虑了算法的探索能力和开发实力，以防止陷入局部最优解的困境。

在 GWO 中， ω 狼更新位置仅受 α 、 β 、 δ 狼的影响，未考虑自身对问题求解的作用。本文保留个体的历史最优解，增强其记忆能力，使其在搜索时始终记住最优解。改进后的位置更新策略如下：

$$\begin{cases} X_{\omega}(t+1) = \frac{X_1(t+1) + X_2(t+1) + X_3(t+1)}{3} \omega + D_{\Sigma} \\ D_{\Sigma} = c_1 r (X_{\alpha}(t) - X_{\omega}(t)) + \\ c_2 r (X_{\beta}(t) - X_{\omega}(t)) + c_3 r (X_{\delta}(t) - X_{\omega}(t)) \\ \omega = (\omega_{\max} - \omega_{\min}) \frac{t}{t_{\max}} + \omega_{\min} \end{cases} \quad (13)$$

式中， ω 为惯性权重； $X(t)$ 表示 t 时刻狼群的位置； c_1 、 c_2 、 c_3 为 α 狼、 β 狼、 δ 狼的学习因子； r 为 $[0, 1]$ 间的随机变量。本文经多次实验取 c_1 为 0.6， c_2 、 c_3 之和为 0.6， ω_{\max} 为 0.9， ω_{\min} 为 0.3。

1.6 IWGO-TCN-BiGRU 模型

本文通过改进后的 IGWO 算法对模型的参数进行寻优，IGWO 算法相较其他算法更不易陷入局部最优，使用 IGWO 优化 TCN-BiGRU 模型的流程如下。

1) 对原始输入数据进行预处理，按照 7:2:1 的比例划分为训练集、验证集、测试集。

2) 参数初始化，利用 Fuch 映射初始化灰狼种群，确定 α 狼、 β 狼、 δ 狼的位置，设置最大迭代次数 t_{\max} ，确定灰狼种群大小 N 。

3) 检查搜索位置是否超出搜索空间范围，若超出，则重新调整搜索位置。

4) 通过训练集数据对模型进行训练，并在验证集上将 RMSE 作为 IGWO 的适应度函数。

5) 开始迭代，计算适应度值并更新，由式 (12) 更新控制系数 a ，并由式 (13) 对灰狼种群的位置进行更新。

6) 计算选举函数 $F(\alpha, \beta)$ 的适应度并迭代更新。

7) 判断是否达到 IGWO 的最大迭代次数，若未达到则继续迭代更新；反之，寻优结束，输出最优的 α 狼，并将其映射为 TCN-BiGRU 的各个参数进行训练。

8) 参数确定完毕，得到最优的 TCN-BiGRU 模型，送入测试集中对负荷数据进行预测。

整体预测流程图如图 6 所示。

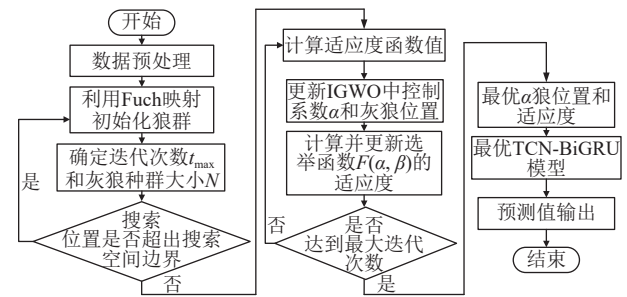


图 6 IWGO-TCN-BiGRU 负荷预测流程图

2 仿真实验及结果分析

本文的实验基于公开的美国某地区住宅负荷数据集^[14]和中国电工数学建模竞赛负荷数据集 (<http://www.csee.org.cn/>)，构建模型进行实验分析。实验输入为过去 10 h 的负荷数据，预测下一时刻的负荷数据。其中，美国住宅负荷数据集记录了 2012 年某地区住宅负荷变化情况，采样间隔为 1 h，包含电力负荷特征以及供暖、室内照明共 12 维特征；中国电工数学建模竞赛负荷数据集则选自地区 1，记录了从 2012 年开始连续 3 年的负荷数

据,采样间隔为15 min,2个数据集均按照7:2:1的比例划分为训练集、验证集和测试集。使用滑窗法对测试集进行测试。

2.1 实验设置

本文实验系统配置为Windows 10操作系统,CPU为Intel Core i7-7700HQ,显卡为NVIDIA GeForce GTX 1050,内存选用16 GB RAM DDR4@2400 MHz,所提出的模型构建在Python (3.9.12)、Tensorflow2.0框架及Keras中运行。

本文实验评价指标包括平均绝对误差(mean absolute error, MAE)、均方根误差(root mean squared error, RMSE)、平均绝对百分比误差(mean absolute percentage error, MAPE)、拟合优度(R square, R^2)、计算量(FLOPs)、参数量(params)。

2.2 子模块改进效果验证

在进行整体模型分析前,为了验证本文改进的TCN结构以及BiGRU结构的有效性,本文在美国住宅负荷数据集上通过控制变量法进行评估,使用MAE、RMSE、MAPE、 R^2 对结果进行评估,对不同残差结构通过不同指标评估的结果如表1所示。

表1 不同残差结构算法效果对比表

残差结构	MAE/kW	RMSE/kW	MAPE/%	R^2
原始残差	0.042	0.047	6.534	0.955
改进残差	0.039	0.045	5.985	0.962

由表1实验结果可知,本文改进的并行池化残差结构有着更小的误差。在此基础上,对本文设计的多尺度卷积结构的作用效果进行分析,实验结果如表2所示。

表2 不同卷积结构算法效果对比表

卷积结构	MAE/kW	RMSE/kW	MAPE/%	R^2
原始卷积	0.039	0.045	5.985	0.962
多尺度卷积	0.036	0.042	5.889	0.965

由表2可以看出,本文在使用多尺度卷积结构的情况下,模型的整体性能进一步得到提升,这也说明了本文多尺度卷积结构的有效性。为了验证ISA模块的有效性,对BiGRU加入ISA模块进行对比分析,实验结果如表3所示。

表3 有无ISA机制算法效果对比表

ISA机制	MAE/kW	RMSE/kW	MAPE/%	R^2
无	0.051	0.076	9.795	0.954
有	0.045	0.065	8.278	0.960

2.3 模型参数寻优

在对整体模型进行效果分析前,首先确定模型

的参数。使用改进的IGWO算法对TCN-BiGRU模型内部常见的一些参数进行寻优。寻优开始时,设置IGWO中的最大迭代次数 t_{max} 为30,灰狼种群大小 N 为30,搜索空间为5。实验选用Adam优化器,对TCN-BiGRU模型参数寻优结果如表4所示。

表4 TCN-BiGRU参数寻优结果

寻优参数	寻优范围	寻优初值	寻优结果
卷积核大小	1~10	(1,2,3)	(1,4,7)
学习率	0.001~0.700	0.05	0.01
迭代次数	1~600	300	500
隐层节点数	2~128	32	24
隐藏层数	1~5	2	3

2.4 实验结果及验证

为了验证本文提出的IGWO-TCN-BiGRU模型具有更好的预测效果,在美国住宅负荷数据集上通过控制变量法验证模型中主要模块的有效性。首先,为了验证本文提出IGWO优化算法的有效性,将其与传统GWO、麻雀搜索算法(sparrow search algorithm, SSA)在相同的数据上进行了比较分析。在划分的验证集上对模型的预测性能进行检验,根据不同评价指标评估的结果如表5所示。

表5 不同优化算法效果对比表

优化算法	MAE/kW	RMSE/kW	MAPE/%	R^2
GWO	0.047	0.051	7.240	0.962
SSA	0.043	0.046	6.603	0.965
IGWO	0.035	0.035	5.049	0.971

从表5中可以看出,改进的IGWO算法相较于传统GWO算法在各项评价指标上均有一定提升,其MAE降低了25.53%,RMSE降低了31.37%,证明了使用IGWO优化算法整体上预测效果相对更好。在使用IGWO作为模型参数优化算法的基础上,对比本文引入的选举机制对实验结果的影响,通过IGWO算法计算 $F(\alpha, \beta)$ 函数的适应度,不断进行迭代直至获得最优解 $[\alpha_{best}, \beta_{best}]$,将其放入选举函数中,从而得到最终的预测推荐值,得到的最终结果为美国住宅数据集 $\alpha_{best}=0.55$, $\beta_{best}=0.45$ 。在中国竞赛数据集上,确定 $\alpha_{best}=0.53$, $\alpha_{best}=0.47$,实验结果如表6所示。

表6 选举机制对实验结果的影响

选举机制	MAE/kW	RMSE/kW	MAPE/%	R^2
无	0.035	0.035	5.049	0.971
有	0.029	0.034	4.974	0.978

由表 6 可以看出, 加入选举机制后, 模型的预测精度在数值上更高, 证明了结合选举机制的改进是可行的。

本文通过消融实验在美国负荷数据集上验证 TCN、BiGRU、ISA、AR 模块的有效性, 实验结果如表 7 所示, 从表 7 中可以看出, 各模块对模型预测精度的提升都发挥着积极作用。在去掉 TCN 和 AR 模块后, 虽然模型的预测准确度有所降低, 但仍然具有良好的预测效果, 这是由于 BiGRU 和 ISA 也能很好地处理序列间的依赖关系。观察消除单一模块对实验结果的影响, 从结果可以看出在去掉 TCN 模块后, 模型整体表现最差, 说明 TCN 结构在整个模型中发挥着重要作用。综合实验结果可以看出, 使用改进的 TCN 可以充分提取数据中的周期信息, 提升短期序列建模能力, ISA 模块可以自适应改变预测任务的权重大小, AR 模块和 BiGRU 模块可以更好地捕获数据间的线性 and 非线性依赖关系。因此模型在短期预测任务中有着更好地表现。

表 7 美国住宅数据集消融实验结果

实验模型	TCN	BiGRU	ISA	AR	MAE/kW	RMSE/kW	MAPE/%	R^2
实验1		II	II		0.045	0.065	8.278	0.960
实验2	II			II	0.039	0.056	6.863	0.965
实验3		II	II	II	0.037	0.053	6.269	0.966
实验4	II		II	II	0.036	0.049	5.612	0.968
实验5	II	II			0.035	0.048	5.403	0.969
实验6	II	II	II		0.032	0.040	5.026	0.972
本文模型	II	II	II	II	0.029	0.034	4.974	0.978

为了验证提出模型的性能, 本文与主流模型进行了实验对比, 为保证实验的公平性, 利用网格搜索法得到其他对比模型的最优参数, 对比的模型包括 DenseNet^[15]、LSTNet-Attn^[16]、TPA-LSTM^[17]、MTGNN^[18]、Autoformer^[19]、FEDformer^[20]。在美国住宅数据集上的对比实验结果如表 8 所示, 为了验证本文改进模型的泛化性能, 采用中国电工数学建模竞赛数据集进行模型对比实验, 实验结果如表 9 所示。

观察表 8 和表 9 可以看出, DenseNet 算法整体表现一般, 相较本文模型 R^2 分别降低了 4.7% 和 6.0%, 本文提出的 IGWO-TCN-BiGRU 模型整体上预测精度最高, 其 MAE 相较于次优的 FEDformer 分别降低了 5.8% 和 3.7%, 其 RMSE 相较 FEDformer 降低了 6.7% 和 7.5%, 这是因为 FEDformer 通过引入频域信息提取技术, 能够捕捉

时间序列中的周期性和隐藏模式, 从而更准确地识别数据内在的规律。

表 8 美国住宅数据集评价指标对比

方法	MAE/kW	RMSE/kW	MAPE/%	R^2
DenseNet	0.052	0.082	9.690	0.932
LSTNet-Attn	0.048	0.075	9.384	0.956
TPA-LSTM	0.043	0.068	8.426	0.963
MTGNN	0.040	0.062	7.995	0.968
Autoformer	0.035	0.053	6.357	0.969
FEDformer	0.031	0.036	5.103	0.972
本文模型	0.029	0.034	4.974	0.978

表 9 中国竞赛数据集评价指标对比

方法	MAE/MW	RMSE/MW	MAPE/%	R^2
DenseNet	215.386	427.685	4.372	0.935
LSTNet-Attn	185.258	383.903	3.237	0.949
TPA-LSTM	173.031	350.506	2.930	0.957
MTGNN	159.881	325.270	2.675	0.969
Autoformer	98.078	171.262	1.807	0.972
FEDformer	83.366	118.171	1.205	0.983
本文模型	80.281	109.308	1.109	0.995

为了评估模型的资源消耗以及部署潜力, 在美国住宅数据集上对模型的计算量以及模型体积进行了测试, 模型计算量统计仅包含推理阶段, 参数优化过程的耗时并未计入, 计算量和参数量如表 10 所示。

表 10 计算量与模型体积对比表

方法	Params/MB	MFLOPs
Autoformer	9.400	126.89
DenseNet	3.700	120
FEDformer	7.300	70.84
MTGNN	0.051	19.38
LSTNet-Attn	0.071	11.41
TPA-LSTM	0.082	1.84
本文模型	0.037	0.64

通过表 10 可以看出, 本文模型在参数量和计算量上均显著低于 DenseNet、Autoformer 和 FEDformer, 同时也低于一些专门设计用于时间序列预测的轻量级模型, 如本文模型与 TPA-LSTM 相比在参数量上降低了一半以上, 同时计算量更低。这表明本文模型在计算量和参数量上的优势使其在资源受限的环境中更具应用潜力。

3 结束语

为了提升短期负荷的预测精度, 本文提出了一种改进 TCN-BiGRU 模型。具体改进如下。

1) 对 TCN 与 BiGRU 模型分别进行改进, 引

入多尺度卷积的TCN能够更好地捕捉电力负荷数据中的局部信息,融合改进ISA机制的BiGRU能够捕捉更多针对电力负荷数据的区间信息。

2) 针对短期电力负荷数据波动性强、外界影响因素大的特点,在组合模型中加入AR模块和本文提出的选举机制提升预测准确率。

3) 针对原始灰狼优化算法容易陷入最优解的问题提出了一种改进灰狼优化算法,改进灰狼优化算法综合考虑了算法的探索能力和局部开发能力,有助于避免算法过早陷入局部最优解,并利用改进算法对本文提出的模型进行参数寻优,增强模型预测能力。

参考文献

- [1] 刘博. 基于 PSO-LSTM 算法的短期电力负荷预测应用研究[D]. 长春: 吉林大学, 2020.
LIU B. Application research of short-term power load forecasting based on PSO-LSTM algorithm[D]. Changchun: Jilin University, 2020.
- [2] 吴晨, 姚菁, 薛贵元, 等. 基于 MMoE 多任务学习和长短期记忆网络的综合能源系统负荷预测[J]. 电力自动化设备, 2022, 42(7): 33-39.
WU C, YAO J, XUE G Y, et al. Load forecasting of integrated energy system based on MMoE multi-task learning and LSTM[J]. Electric Power Automation Equipment, 2022, 42(7): 33-39.
- [3] LIN B D, ZHANG X Y, ZHANG M, et al. Improved genetic algorithm-based research on optimization of least square support vector machines: An application of load forecasting[J]. *Soft Computing*, 2021, 25(18): 11997-12005.
- [4] ALOTAIBI M A. Machine learning approach for short-term load forecasting using deep neural network[J]. *Energies*, 2022, 15(17): 6261.
- [5] CHAN J W, YEO C K. A Transformer based approach to electricity load forecasting[J]. *The Electricity Journal*, 2024, 37(2): 107370.
- [6] ZHOU Y, WANG L Y, QIAN J H. Application of combined models based on empirical mode decomposition, deep learning, and autoregressive integrated moving average model for short-term heating load predictions[J]. *Sustainability*, 2022, 14(12): 7349.
- [7] 孙隽丰, 李成海, 曹波. 基于 TCN-BiLSTM 的网络安全态势预测[J]. 系统工程与电子技术, 2023, 45(11): 3671-3679.
SUN J F, LI C H, CAO B. Network security situation prediction based on TCN-BiLSTM[J]. Systems Engineering and Electronics, 2023, 45(11): 3671-3679.
- [8] BAI S J, KOLTER J Z, KOLTUN V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling[EB/OL]. [2024-04-22]. <https://arxiv.org/abs/1803.01271>.
- [9] ZHANG D X, WANG S, LIANG Y Q, et al. A novel combined model for probabilistic load forecasting based on deep learning and improved optimizer[J]. *Energy*, 2023, 264: 126172.
- [10] 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.
- [11] PU X W, XIAO H, WANG J R, et al. A novel GRU-TCN network based interactive behavior learning of multi-energy microgrid under incomplete information[J]. *Energy Reports*, 2023, 9: 608-616.
- [12] HE K, SU Z Q, TIAN X Q, et al. RUL prediction of wind turbine gearbox bearings based on self-calibration temporal convolutional network[J]. *IEEE Transactions on Instrumentation and Measurement*, 2022, 71: 3501912.
- [13] 余德荧, 李厚朴, 纪兵, 等. 基于灰狼优化算法的快速选星方法[J]. 系统工程与电子技术, 2023, 45(5): 1489-1495.
YU D Y, LI H P, JI B, et al. Fast satellite selection method based on grey wolf optimization algorithm[J]. *Systems Engineering and Electronics*, 2023, 45(5): 1489-1495.
- [14] Open energy data initiative (OEDI)[EB/OL]. [2024-06-25]. <https://data.openei.org/submissions/153>.
- [15] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 2261-2269.
- [16] LAI G K, CHANG W C, YANG Y M, et al. Modeling long- and short-term temporal patterns with deep neural networks[C]// Proceedings of the 41st International ACM SIGIR Conference on Research & Development in Information Retrieval. New York: ACM, 2018: 95-104.
- [17] SHIH S Y, SUN F K, LEE H Y. Temporal pattern attention for multivariate time series forecasting[J]. *Machine Learning*, 2019, 108(8): 1421-1441.
- [18] WU Z H, PAN S R, LONG G D, et al. Connecting the dots: Multivariate time series forecasting with graph neural networks[C]// Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. New York: ACM, 2020: 753-763.
- [19] WU H X, XU J H, WANG J M, et al. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting[EB/OL]. (2021-06-24)[2024-10-11]. <https://arxiv.org/abs/2106.13008>.
- [20] ZHOU T, MA Z Q, WEN Q S, et al. FEDformer: frequency enhanced decomposed transformer for long-term series forecasting[C]// Proceedings of the 39th International Conference on Machine Learning. Baltimore: [s.n.], 2022: 27268-27286.