光伏运维中融合KAN Transformer的TCN深度 学习模型及故障检测

马 寅,张 鹏,张宪康,李 灏,姬艳秋,张炜明,赵进国,张旭军

国网甘肃省电力公司,甘肃 兰州

收稿日期: 2024年10月2日; 录用日期: 2024年11月20日; 发布日期: 2024年11月29日

摘要

光伏发电作为清洁能源的重要组成部分,在电力系统中的集成对提高能源利用效率和减少环境污染具有 重要意义。但光伏发电的波动性和不确定性给电网的稳定运行带来了挑战。本研究针对甘肃地区电力系 统中光伏发电的预测问题,提出了一种创新的深度学习模型,该模型融合了Kolmogorov-Arnold Networks (KAN)、Transformer和时间卷积网络(TCN)的优势,以提高预测精度。考虑到光伏发电数据的高 维性和非线性特征,KAN层被引入以提取复杂的数据模式。Transformer层通过自注意力机制有效地捕 捉了时间序列中的长距离依赖关系。TCN层利用扩张卷积技术进一步增强了模型对局部时间特征的捕捉 能力。为了处理数据集中的不平衡问题,本研究采用了SMOTE技术进行数据预处理,以增强模型对少数 类别的识别能力。在模型训练过程中,均方误差(MSE)损失函数被用作优化目标,以最小化预测误差。此 外,本研究不仅关注预测精度,还将模型应用于光伏发电故障监测,通过实时监控和分析运行数据,实 现了故障的早期识别和预警。实验结果表明,所提出的模型在光伏发电预测任务上表现出色,相比于现 有方法具有更高的预测精度和鲁棒性。此外,通过故障监测的应用案例,证明了模型在实际电力系统运 维中的实用性。本研究的成果为光伏发电的预测和运维管理提供了新的视角和技术支持,对促进可再生 能源的高效利用具有重要意义。

关键词

光伏发电预测,Transformer,时间卷积网络,故障监测

TCN Deep Learning Model and Fault Detection Based on Integrating KAN Transformer in Photovoltaic Operation and Maintenance

Yin Ma, Peng Zhang, Xiankang Zhang, Hao Li, Yanqiu Ji, Weiming Zhang, Jinguo Zhao, Xujun Zhang

文章引用:马寅,张鹏,张宪康,李灏,姬艳秋,张炜明,赵进国,张旭军.光伏运维中融合 KAN Transformer 的 TCN 深度学习模型及故障检测[J].人工智能与机器人研究, 2024, 13(4): 947-960. DOI: 10.12677/airr.2024.134096

State Grid Gansu Electric Power Company, Lanzhou Gansu

Received: Oct. 2nd, 2024; accepted: Nov. 20th, 2024; published: Nov. 29th, 2024

Abstract

Photovoltaic (PV) power generation, as a vital component of clean energy, plays a significant role in integrating into the power system, enhancing energy utilization efficiency, and reducing environmental pollution. However, the volatility and uncertainty of photovoltaic power generation pose challenges to the stable operation of the power grid. This study addresses the forecasting issue of photovoltaic power generation in the Gansu region's power system by proposing an innovative deep learning model that integrates the advantages of Kolmogorov-Arnold Networks (KAN), Transformers, and Temporal Convolutional Networks (TCN) to improve forecasting accuracy. Considering the high-dimensional and nonlinear characteristics of photovoltaic power generation data, KAN layers are introduced to extract complex data patterns. Transformer layers effectively capture long-range dependencies in time series through self-attention mechanisms. TCN layers further enhance the model's ability to capture local temporal features using dilated convolution techniques. To address the imbalance in the dataset, this study employs the SMOTE technique for data preprocessing, enhancing the model's recognition capability for minority classes. During model training, the mean squared error (MSE) loss function is used as the optimization objective to minimize prediction errors. Moreover, this study not only focuses on forecasting accuracy but also applies the model to photovoltaic power generation fault monitoring, achieving early identification and warning of faults through real-time monitoring and analysis of operational data. Experimental results show that the proposed model performs excellently in photovoltaic power generation forecasting tasks, with higher prediction accuracy and robustness compared to existing methods. Additionally, through the application of fault monitoring, the model's practical value in actual power system operations is demonstrated. The outcomes of this study provide new perspectives and technical support for the forecasting and operational management of photovoltaic power generation, which is of great significance for promoting the efficient use of renewable energy.

Keywords

Photovoltaic Power Generation Forecasting, Transformer, Temporal Convolutional Network, Fault Monitoring

Copyright © 2024 by author(s) and Hans Publishers Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). http://creativecommons.org/licenses/by/4.0/

CC O Open Access

1. 引言

甘肃省作为我国重要的能源基地,特别是在太阳能发电领域,近年来取得了显著的发展。根据 2024 年 6 月 13 日的资料,甘肃省 5 月份光伏发电量达到 32.28 亿千瓦时,同比增长 75.35%,占发电总量的 18.32%。截至 5 月底,全省发电装机容量为 9462.62 万千瓦,其中太阳能装机容量为 2744.02 万千瓦,同 比增长 67.35% [1]。在光伏系统运维方面,随着大数据、云计算、5G 等技术的发展,智能化电站建设布 局逐步完善,可以主动对电站运维过程中出现的问题进行预警提醒,光伏电站智能化运维发展主要体现 在组件智能清扫机器人、智能运维无人机系统等应用领域[2]。 甘肃省在新能源发展上具有得天独厚的优势,风能和光能资源有效储量分别位居全国第五和第三, 具备基地化、规模化、一体化开发的优越条件。此外,甘肃省在"十四五"期间将加快推进风电基地建设 和持续扩大光伏发电规模,特别是在沙漠、戈壁、荒漠地区开展规模化建设,实现太阳能发电与荒漠化 治理、生态修复、农牧业融合发展[3]。

随着可再生能源的迅速发展,光伏发电作为其中的重要组成部分,在电力系统中扮演着越来越重要的角色。甘肃凭借其丰富的太阳能资源,已成为光伏发电潜力巨大的地区之一。然而,光伏发电的间歇 性和不稳定性对电网的稳定运行提出了挑战,如何准确预测光伏发电量对甘肃电力系统的优化调度、负 荷平衡和能源管理至关重要[4]。

近年来,深度学习在各类时序数据处理任务中展现了强大的能力,尤其是在复杂、高维数据的建模和分类任务上。传统的神经网络,如卷积神经网络(CNN)和循环神经网络(RNN),已经被广泛应用于时序数据处理。虽然 CNN 在图像处理和时序数据的局部特征提取方面表现出色,但其难以捕捉长程依赖关系。此外,CNN 对高维数据的处理能力也有限,难以有效处理复杂的多维时序数据[5]。

循环神经网络(RNN)及其变体(如 LSTM、GRU)擅长处理时序数据中的时间依赖性,然而,其在捕捉 长程依赖关系时存在梯度消失或爆炸的问题。尽管 LSTM 和 GRU 通过门控机制部分缓解了这些问题, 但它们的训练效率较低,并且在处理长序列时性能会下降[6]。

一些研究者尝试将 CNN 和 RNN 结合,以捕捉局部和全局的时序模式。然而,这类混合模型在处理 高维数据和复杂时序依赖关系时,仍然面临计算复杂度和模型效率的挑战。为了应对这一挑战,本研究 提出了一种融合了 Kolmogorov-Arnold Networks (KAN)、Transformer 和时间卷积网络(TCN)的深度学习 模型。KAN 通过函数分解,将复杂的高维时序数据映射到低维空间,从而降低数据的复杂度,并为后续 的 Transformer 和 TCN 提供更有效的特征表示; Transformer 通过自注意力机制捕捉全局依赖关系,而 TCN 则通过卷积操作捕捉局部特征。两者结合能够有效处理数据中的全局和同步模式。TCN 的并行处理 能力和 Transformer 的自注意力机制使得该融合模型在处理长序列时具备高效计算能力,同时能够捕捉重 要的长程依赖关系。将 Kolmogorov-Arnold Networks (KAN)、Transformer 和时间卷积网络(TCN)进行融 合,是解决复杂时序分类任务的一种有效方法。相比传统的神经网络模型,这种融合策略结合了 KAN 在 高维数据处理上的优势、Transformer 在长程依赖关系捕捉上的优势以及 TCN 在局部特征提取和高效计 算上的优势[7]-[9]。因此,本文的融合模型在处理复杂时序数据时,展现出更优的性能和更强的适应性。

考虑到光伏发电量受到多种因素的影响,如气象条件(太阳辐射、温度、湿度等)、地理位置、时间特征等。在进行光伏发电量预测前,特征选择是数据处理的关键步骤,它可以有效提高模型的预测性能,减少计算复杂度,并防止过拟合。我们在此采用了 Synthetic Minority Over-sampling Technique (SMOTE) 进行数据预处理,SMOTE 通过生成合成样本来平衡各类别的比例,从而提高模型对少数类别的识别能力。在模型训练阶段,我们采用了均方误差(Mean Squared Error, MSE)作为损失函数来衡量预测值与实际观测值之间的差异。MSE 损失函数的优化有助于模型学习到更为精确的预测参数[10]。

此外,本研究还扩展了模型的应用范围,将其用于光伏发电故障监测。通过对光伏系统运行数据的 实时监控和分析,模型能够及时识别并预警潜在的故障,为电力系统的稳定运行提供保障。

总之,本研究的光伏发电预测模型不仅提高了预测精度,还通过数据预处理和故障监测增强了模型 的实用性和可靠性。这对于甘肃乃至更广泛地区的电力系统运维具有重要的参考价值。

2. 数据收集

本研究旨在建立一个神经网络模型,利用甘肃 2024 年 1 月 1 日至 2024 年 6 月 30 日期间太阳能发电 量实际数据预测未来的发电量数据。表 1 显示观测值总数(count)为 17,362 个,平均值(Mean)为 3823.52, 表明数据集的中心位置;数据的分散程度由标准差(Standard Deviation) 4862.58 体现,显示出观测值在平 均值周围的分布范围。数据的分布情况通过百分位数来进一步阐释,其中第25百分位数(O1)为8.78,中 位数(Median)为 214.18, 而第 75 百分位数(Q3)为 8206.70。最小值和最大值分别为 0 和 17467.30, 揭示了 数据的极端值范围。这些统计量为理解数据的一般趋势和分布特性提供了重要信息。

Table 1. Descriptive statistical tables 表 1. 描述性统计表

输入变	样本量	平均值	标准差	最小值	第 25 位百	第 50 位百	第 75 位百	最大值
量	(个)	(MW)	(MW)	(MW)	分数(MW)	分数(MW)	分数(MW)	(MW)
	17,362	3823.51	4862.57	0	8.78	214.17	8206.7	17467.3

本研究利用 Python 编程语言导入了必要的库来进行数据可视化。通过绘制预测光伏发电量随时间的 变化趋势图,我们不仅直观地记录了光伏发电预测数据的动态行为,还揭示了它与太阳能发电量的潜在 相关性。数据可视化的结果为模型的建立提供了直观的依据,有助于识别和分析太阳能发电量真实数据 和预测数据之间的复杂关系,从而为太阳能发电量的预测和优化提供了科学依据。

数据预处理

模型的开发分为两个关键阶段:数据预处理和模型训练。预处理阶段的关键方面是涉及解决数据集 中固有的不平衡问题。原始数据分布如图1所示。图1描绘了"正常"和"异常"类别的原始分布,显 示出与"正常"类别的显著差异,大大超过了"异常"类别的数量。这种不平衡可能会在模型中引入偏 差,导致泛化性能不佳,特别是对于少数群体而言。对此本文引入 SMOTE 技术来缓解这个问题。这种 技术产生了少数群体的合成例子,旨在平衡阶级分布而不丢失有价值的信息。



正常和异常数据的分布

Figure 1. The raw distribution of normal and outlier values 图 1. 正常值和异常值的原始分布

SMOTE 过采样方法主要是为了解决数据集中类别不平衡的问题,尤其是少数类样本数量过少的情况。其基本原理是对少数类样本进行人工合成来增加样本数量。它通过在少数类样本与其最近邻的少数 类样本之间进行线性插值来生成新的样本[11]。

SMOTE 的计算公式:假设少数样本中的一个样本点为x,其最近邻的少数类样本为x_m。

$$x_{new} = x + rand(0,1) \times (x_{nn} - x)$$
⁽¹⁾

其中 x_{new} 为生成的新样本, rand (0,1) 是一个 0 到 1 之间的随机数。

SMOTE 操作的结果如图 2 所示。该图展示了数据集经过 SMOTE 操作后的分布,其中"正常"和 "异常"类的频率几乎相同,表明成功平衡了数据集。而准备一个平衡的数据集对于训练稳健的异常检 测模型至关重要,确保模型不会过度偏向多数类。这一基础预处理步骤有助于更公平、更准确地评估模 型在所有类别中的性能,这在异常检测中至关重要。因为在异常检测中,错误分类异常的成本可能很高。



正常和异常数据的分布

Figure 2. Distribution of normal and abnormal values after SMOTE 图 2. SMOTE 后正常值和异常值的分布

3. 特征选择

在学习模型的开发过程中,利用特征选择的方法来限制输入变量的数量,从而减少计算量。在将数据输入机器学习系统时,特征选择可以消除不必要或重复的信息。输入更多的历史数据和输入变量可以较好地提取数据特征,但这样会增加模型的复杂性。因此,确定输入参数的数量至关重要。皮尔逊相关系数(PCC)用于评估影响光伏发电量关键输入特性。PCC 经常用于衡量人工智能中特征选择的两个随机变量之间的线性依赖性。其计算形式如下:

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \overline{x}) \sum_{i=1}^{n} (y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \overline{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \overline{y})^2}}$$
(2)

公式(2)中 r_{xy} 显示x与y有线性相关性,并且 $|r_{xy}| \le 1$ 。当 $r_{xy} = 1$ 时, x与y呈正相关; 当 $r_{xy} = 0$ 时, x与y没有线性关; 当 $r_{xy} = -1$ 时, x与y呈负相关。在模型的学习阶段,根据大于或等于 0.5 的阈值确定可接受和排除的参数。特征选择的关键在于识别和选择与目标变量最相关的特征,以提高模型的性能和泛化能力。本模型中,特征选择通过命令行参数实现,使得模型训练可以灵活地适应不同的数据特征。

4. 模型建立

4.1. 损失函数

损失函数(Loss Function)用于衡量模型预测值与实际观测值差异的程度,指导模型训练过程中的参数 优化。本文采用均方误差(Mean Squared Error, MSE)损失函数。它是一种连续且可微的函数,这一特性使 得它在基于梯度的优化算法中表现出色。在随机梯度下降(Stochastic Gradient Descent, SGD)及其变体的优 化方法中,MSE 的梯度容易计算,从而能够有效地引导模型参数的更新,加速模型的训练收敛过程。

由于 MSE 对误差进行了平方运算,这使得较大的误差被赋予更高的权重。在实际应用中,对于那些 偏离真实值较大的预测, MSE 能够给予更强烈的惩罚,促使模型更加注重减少这些较大的偏差,从而提 高整体的预测精度。通过对 MSE 的分析,可以深入理解模型的性能和行为,为模型的改进和优化提供有 力的理论支持[10]。其性能和效果在大量的研究和实际应用中得到了充分的验证和认可。均方误差 MSE 的计算公式如下:

$$MSE = \frac{1}{m} \sum_{i=1}^{m} (y_i - f_i)^2$$
(3)

其中 y_i 指实际值, f_i 代表模型的预测输出值, m为样本容量。通过先进的优化算法(如高效的 Adam 优化器[12])持续不断地调整模型的参数,以最大程度地减小损失函数的值,从而促使模型能够精准无误 地预测未来的时间序列数值,为实际应用提供可靠的决策支持[8]。

4.2. Kolmogorov-Arnold (KAN)

KAN 以 Kolmogorov-Arnold 表示定理为基石,其核心原理在于通过多层级的线性与非线性组合来近 似复杂的函数形态。对于给定的输入向量 $x = (x_1, x_2, \dots, x_n)$, KAN 的某一层可以精确地表述为:

$$f(x) = g\left(\sum_{i=1}^{k} p_i g_p(x_i)\right)$$
(4)

在此表达式中, *p_i*和*g_p*是可学习的参数, *x_i*是输入*x*的变换。*k*是定理中的常数。而激活函数*g*通常采用诸如 ReLU、Sigmoid 或 Tanh 等非线性函数,为模型赋予强大的非线性拟合能力,使得模型能够捕捉到数据中丰富多样的复杂模式和非线性关系[5]。图 3 展示了 KAN 模型的简单结构示意图,该图清晰地描绘了在边缘上的可学习激活函数和在节点上的求和操作,这两个关键组件共同构成了 KANs 的核心架构。

4.3. Transformer 模型

Transformer 是一种用于自然语言处理(NLP)和其他序列到序列(sequence-to-sequence)任务的深度学习 模型架构,它在 2017 年由 Vaswani 等人首次提出。以下是 Transformer 的一些重要组成部分和特点:

自注意力机制(Self-Attention): 这是 Transformer 的核心概念之一,它使模型能够同时考虑输入序列中的所有位置,而不是像循环神经网络(RNN)或卷积神经网络(CNN)一样逐步处理。自注意力机制允许模型根据输入序列中的不同部分来赋予不同的注意权重,从而更好地捕捉语义关系。



Figure 3. A schematic diagram of the KAN Model's structure 图 3. KAN 模型的简单结构示意图

多头注意力(Multi-Head Attention): Transformer 中的自注意力机制被扩展为多个注意力头,每个头可以学习不同的注意权重,以更好地捕捉不同类型的关系。多头注意力允许模型并行处理不同的信息子空间。

堆叠层(Stacked Layers): Transformer 通常由多个相同的编码器和解码器层堆叠而成。这些堆叠的层 有助于模型学习复杂的特征表示和语义。

位置编码(Positional Encoding):由于 Transformer 没有内置的序列位置信息,它需要额外的位置编码 来表达输入序列中单词的位置顺序。

残差连接和层归一化(Residual Connections and Layer Normalization):这些技术有助于减轻训练过程中的梯度消失和爆炸问题,使模型更容易训练。

编码器和解码器: Transformer 通常包括一个编码器用于处理输入序列和一个解码器用于生成输出序列,这使其适用于序列到序列的任务,如机器翻译等[8]。

Transformer 模型的核心精髓在于其独特的自注意力机制(Self-Attention)。对于输入序列 $X = (x_1, x_2, \dots, x_n)$, 其中每个元素 $x_i \in \mathbb{R}^d$, 注意力机制的计算过程如下:

Attention
$$(Q, K, V) = \text{Softmax}\left(\frac{QK^{T}}{\sqrt{d_{k}}}\right)V$$
 (5)

式中 $Q = XW_Q$, $K = XW_K$, $V = XW_V$ 。 W_Q , W_K , W_V 均为可学习的参数矩阵, 它们在训练过程中不断优化, 以捕捉输入序列中元素之间的内在联系和依赖关系。而 $d_k \ge K$ 的维度, 通过除以 $\sqrt{d_k}$ 进行缩放, 有助于稳定训练和优化注意力得分的分布。

通过这种自注意力机制,模型能够动态地为输入序列中的每个元素分配不同的权重,从而聚焦于对 当前预测任务最为关键的部分,有效地捕捉长序列中的各种特征依赖关系。Transformer 模型图如图 4 所 示。

图 4 清晰地展示了 Transformer 模型关键组成部分。输入首先经过嵌入层进行编码,然后依次通过编码器和解码器模块。编码器中,多边注意力机制使模型能够关注输入序列的不同部分,经过 Add&Norm 层进行归一化和残差连接。解码器同样包含多边注意力和 Add&Norm 层,同时还有掩码多头注意力用于



处理特定的顺序关系。最终通过线性层和 Softmax 输出预测结果。

Figure 4. Transformer model 图 4. Transformer 模型

4.4. 时间卷积网络(TCN)

TCN (Temporal Convolutional Network)充分利用一维卷积操作来处理时间序列数据。假设输入序列为 $x = (x_1, x_2, \dots, x_n)$,卷积核大小为k,步长为s,则卷积运算可以精准地表示为:

$$y_{i} = \sum_{i=0}^{k-1} w_{j} x_{i+j \times s}$$
(6)

其中 w_j 是卷积核的权重,其数值决定了卷积操作对输入数据的滤波效果。通过调整卷积核的大小、数量和步长等参数,TCN 能够有效地提取时间序列数据中的局部特征和短期模式[9]。

TCN结构图如图5所示,图中展示了内核为三,扩张卷积为[1,2,4,8]的膨胀非因果卷积构成的TCN。

 扩张卷积为4,即上一层每前推4个时间步的输出,作为这一层的输入。
 输出 扩张卷积-8
 除藏层 扩张卷积-4
 隐藏层 扩张卷积-4
 隐藏层 扩张卷积-2
 隐藏层 扩张卷积-2
 隐藏层 扩张卷积-1
 隐藏层
 扩张卷积-2
 隐藏层
 扩张卷积-1
 隐藏层
 扩张卷积-2
 隐藏层
 扩张卷积-1
 输入

图 5. TCN 结构

本文通过构建融合模型首先将输入的时间序列数据经由嵌入层转换为低维向量表示,从而降低数据 的维度并提取初始的特征信息。随后,数据通过 KAN 层进行初步的特征提取和非线性变换,为后续的处 理奠定基础。紧接着,利用 Transformer 的自注意力机制全面捕捉长序列中的依赖关系,充分获取全局上 下文信息。最后,通过 TCN 层进一步精细地提取局部时间特征,从而实现对时间序列数据的全面且深入 的理解和分析。

在研究中,本文所构建的深度学习模型吸取了卷积神经网络(CNN)和循环神经网络(RNN)的长处,巧 妙地融合了 Kolmogorov-Arnold Networks (KAN)、Transformer 和时间卷积网络(TCN)的独特优势,旨在 高效处理时间序列数据并实现精准预测。

5. 基于融合模型的实证结果分析

此处我们使用了 2024 年 1 月 1 日至 2024 年 6 月 30 日期间甘肃的历史光伏发电量数据(每日 96 个时段)在神经网络模型中进行训练,结果如图 6 所示。在模型训练过程中,真实值(蓝色曲线)与预测值(红色曲线)几乎重合,表明模型模拟效果很好,我们会通过模型得到良好的预测值。



Figure 6. Model training true and predicted line charts 图 6. 模型训练真实值和预测值折线图

同时,在训练迭代时得到了如图 7 所示的损失值的折线图。该图展示了模型训练过程中损失函数值 随迭代次数(Epoch)的变化情况。横轴表示训练迭代的次数,纵轴表示损失函数的值,范围从 0.0 至 0.6, 以 0.1 为间隔递增。图中的折线通过数据点平滑地描绘了损失值的下降趋势。在初始阶段,损失函数值较 高,随着迭代次数的增加,损失值显著下降,表明模型在训练集上的性能逐渐提升。数据点(0,0.6)表示初 始损失值,而点(175,0.1)则表示经过 175 次迭代后达到的最低损失值。



Figure 7. A line chart where the value of the loss function changes with the number of iterations 图 7. 损失函数值随迭代次数变化的折线图

本文在训练过程中还得到了如图 8 所示的决定系数曲线-*R*² 曲线。该曲线图表示了决定系数随训练轮 数(Epoch)的变化情况。决定系数 *R*²用于衡量回归模型对数据的拟合程度, *R*²的取值范围在 0 到 1 之间。 在训练过程中观察 *R*²随 Epoch 的变化,可以了解模型在不同训练阶段对数据的拟合能力的提升或下降情 况。由图可见,随着训练轮数 Epoch 的增加, *R*² 逐渐增大并趋近于 1,这表明模型的性能在不断改善, 对数据的拟合效果越来越好,拟合结果和预测数据会更优良。

最后,本文基于融合 Kolmogorov-Arnold Networks (KAN)、Transformer 和时间卷积网络(TCN)的优势 所建立的深度学习模型得到了甘肃光伏发电 110 个预测值,并与真实的发电量进行对比,如图 9 所示。 该图展示了光伏发电系统在连续 110 个时间单位内的实际发电量与预测发电量之间的对比。通过折线图 的形式,我们可以直观地观察到两个数据序列随时间变化的趋势和差异。图中的两条折线分别代表了甘 肃光伏实际发电量(真实值)和模型预测发电量(预测值),两者均以兆瓦(MW)为单位。X 轴表示数据量,即 时间单位的序列,而 Y 轴表示光伏发电量,范围从 0 至 14,000 兆瓦。

从图 9 中可以明显看出,在大部分时间单位内,预测值与实际值的变化趋势保持一致,表明所采用 的预测模型能够较好地捕捉光伏发电系统的实际发电行为。然而,也存在若干时间点,在这些点上预测 值与实际值之间出现了偏差,这种偏差往往是由诸如天气变化、系统维护或设备性能波动等多种因素引 起的。特别是在数据量为 40 至 60 的时间单位范围内,预测值与实际值之间的偏差较为显著,这提示我 们模型在这些时段的预测精度有待提高,需要进一步分析导致这些偏差的原因,并探索改进模型预测精 度的方法,这将是未来工作的重要方向。

马寅 等



Figure 9. Line charts of true and predicted values 图 9. 真实值和预测值的折线图

总体而言,图 9 为评估和理解光伏发电预测模型的性能提供了直观的视角,这一良好的效果有力地 证明了所采用模型在数据处理和模式识别方面的出色能力,以及在预测任务中的高度准确性和可靠性, 为相关领域的研究和应用提供了极具价值的参考和支持。

6. 光伏发电故障检测方法

6.1. 数据之间的 Pearson 相关系数

为了建立故障检测方法,本文首先将一天的光伏发电量(光伏发电量的实际值和预测值)和此前七天的实际光伏发电量的相关性进行了计算,得到了表 2 所示的 Pearson 相关系数。

 Table 2. The Pearson correlation coefficient between the daily photovoltaic electricity generation and the actual electricity generation of the preceding seven days

 表 2. 光伏一天发电量和前七天实际发电量的 Pearson 相关系数

变量	S1	S2	S3	S4	S5	S6	S7
实际值	0.98	0.98	0.98	0.98	0.98	0.98	0.98
预测值	0.98	0.98	0.98	0.97	0.96	0.97	0.97

从表 2 中可以看出,一天实际发电量和前七天实际发电量(S1, S2, S3, S4, S5, S6,S 7 表示此前七天的 实际光伏发电量)之间的相关系数都是 0.98,说明它们的相关性非常强;而一天预测发电量和前七天的实 际发电量相关系数大都是 0.97 和 0.98,则证明我们的预测发电量和七个变量的相关性强,并且预测发电 量相对准确。如果我们的预测发电量和七个变量中的某些相关性较弱,那么证明我们的预测数据有问题, 此时则应该在模型中寻找问题并予以解决。

6.2. ROC 曲线原理

为了证明我们的模型的优良性和更好地对预测模型进行故障检测,本文又引入了 ROC 曲线和 AUC 值。ROC (Receiver Operating Characteristic)曲线用于评估分类模型在不同决策阈值下的表现。它通过描绘 真正率(True Positive Rate, TPR)和假正率(False Positive Rate, FPR)的关系来衡量模型的性能。

真正率(TPR):即召回率,表示在所有实际为正类的样本中被正确预测为正类的比例,计算公式为:

$$TPR = \frac{TP}{TP + FN}$$
(7)

其中 TP (True Positive)是将正类预测为正类的样本数, FN (False Negative)是将正类预测为负类的样本数。

假正率(FPR): 表示在所有实际为负类的样本中被错误预测为正类的比例, 计算公式为:

$$FPR = \frac{FP}{FP + TN}$$
(8)

其中 FP (False Positive)是将负类预测为正类的样本数, TN (True Negative)是将负类预测为负类的样本数。

6.3. AUC 值的原理

AUC (Area Under the Curve)是 ROC 曲线下的面积,表示模型区分正负样本能力的大小。AUC 的值 介于 0.5 到 1 之间,值越大表示模型的分类性能越好。AUC = 0.5 表示模型的分类能力相当于随机猜测; AUC = 1 表示模型的分类能力完美,即对正负类样本完全正确分类[11]。

AUC 值的数学计算可以通过积分 ROC 曲线下的面积来实现。通常使用梯形规则(Trapezoidal Rule)近 似计算 AUC,即将 ROC 曲线分为多个小梯形,并求和各个梯形的面积。ROC 曲线下的面积 AUC 可通 过如下积分表示:

$$AUC = \int_{0}^{1} TPR(FPR) d(FPR)$$
(9)

根据上式由光伏发电预测值计算得到 ROC 曲线和 AUC 值,如图 10 所示。图中 ROC 曲线(黄色折 线)下面积(Area Under the Curve, AUC)为 0.72,该值量化了模型在区分正负样本时的整体效果。AUC 值 越接近 1.0,表明模型的分类性能越优。本研究所得的 ROC 曲线清晰地位于随机猜测线之上,表明所提 出的故障检测模型整体效果不错。



Figure 10. ROC curve and AUC value 图 10. ROC 曲线和 AUC 值

7. 结束语

本研究提出的深度学习模型,通过融合 Kolmogorov-Arnold Networks (KAN)、Transformer 和时间卷 积网络(TCN)的技术优势,成功解决了光伏发电预测中的高维性和非线性问题,提高了预测精度。在技术 应用层面,模型的创新之处在于其对时间序列长距离依赖关系的捕捉能力,以及对局部时间特征的强化, 这为光伏发电数据的复杂模式识别提供了新的视角。在数据预处理上,本文采用 SMOTE 技术对数据集 进行预处理,有效解决了不平衡问题,增强了模型对少数类别的识别能力,这在提高模型泛化能力方面 起到了关键作用。对优化目标,本文使用均方误差(MSE)作为损失函数,确保了模型训练过程中预测误差 的最小化,这为模型的精确度提供了量化的优化标准。就实际应用价值而言,模型不仅在光伏发电预测 任务上表现出色,而且通过实时监控和分析运行数据,实现了故障的早期识别和预警,证明了其在实际 电力系统运维中的实用价值。在社会与环境影响方面,本文的研究成果对于提高能源利用效率、减少环 境污染以及促进可再生能源的高效利用具有重要意义,可为电力系统的可持续发展提供良好的技术支持。 尽管本文取得了有意义的研究成果,但光伏发电预测和故障监测领域仍存在进一步优化和研究的空间, 未来的工作可以探索模型在不同环境条件下的适应性,以及与其他预测技术的集成应用。

基金项目

本文由国网甘肃省电力公司科技项目"基于长周期运行的双边现货市场全链条智能技术支持及运维 关键技术研究及应用"(522722240008)资助。

参考文献

- [1] 乐天. 甘肃省源荷分布不平衡问题及对策研究[D]: [硕士学位论文]. 北京: 华北电力大学, 2022.
- [2] 张华,李明. 大数据与云计算在光伏系统运维中的应用研究[J]. 可再生能源, 2024, 34(4): 50-58.
- [3] 甘肃省发展和改革委员会. 甘肃省"十四五"能源发展规划[R]. 兰州: 甘肃省发展和改革委员会, 2021.
- [4] 汪宁渤,马明,强同波,等.高比例新能源电力系统的发展机遇、挑战及对策[J].中国电力,2018,51(1):29-35,50.
- [5] King, M., Woo, S.I. and Yune, C. (2024) Utilizing a CNN-RNN Machine Learning Approach for Forecasting Time-Series Outlet Fluid Temperature Monitoring by Long-Term Operation of BHEs System. *Geothermics*, 122, Article ID: 103082. <u>https://doi.org/10.1016/j.geothermics.2024.103082</u>
- [6] Kisvari, A., Lin, Z. and Liu, X. (2021) Wind Power Forecasting—A Data-Driven Method along with Gated Recurrent Neural Network. *Renewable Energy*, 163, 1895-1909. <u>https://doi.org/10.1016/j.renene.2020.10.119</u>
- [7] Liu, Z., Wang, Y., Vaidya, S., Ruehle, F., Halverson, J., Söljčić, M., Hou, T.Y. and Tegmark, M. (2024) KAN: Kolmogorov-Arnold Networks. arXiv: 2404.19756.
- [8] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Polosukhin, I., *et al.* (2017) Attention Is All You Need. arXiv: 1706.03762.
- Bai, S., Kolter, J.Z. and Koltun, V. (2018) An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling arXiv: 1803.01271.
- [10] Wang, Y. and Wang, Y. (2020) A Hybrid Machine Learning Model for Time Series Prediction. *Information Fusion*, 52, 81-90.
- [11] Dablain, D., Krawczyk, B. and Chawla, N.V. (2023) Deepsmote: Fusing Deep Learning and SMOTE for Imbalanced Data. IEEE Transactions on Neural Networks and Learning Systems, 34, 6390-6404. <u>https://doi.org/10.1109/tnnls.2021.3136503</u>
- [12] Kingma, D.P. and Ba, J. (2015) Adam: A Method for Stochastic Optimization. arXiv: 1412.6980.