

基于 KLPP-K-means-BiLSTM 的台区短期 电力负荷预测

朱 江^{1*}, 汪 帆¹, 曹春堂¹, 易灵芝², 邹嘉乐³

(1. 中车时代电动汽车股份有限公司, 湖南 株洲 412000;

2. 湘潭大学 自动化与电子信息学院, 湖南 湘潭 411100;

3. 湖南交通工程学院 交通运输工程学院, 湖南 衡阳 421001)

KLPP-K-means-BiLSTM Based Short-Term Power Load Forecasting for Station Areas

ZHU Jiang^{1*}, WANG Fan¹, CAO Chuntang¹, YI Lingzhi², ZOU Jiale³

(1. CRRC Electric Vehicle Co., Ltd., Zhuzhou 412000, China;

2. School of Automation and Electronic Information, Xiangtan University, Xiangtan 411100, China;

3. Faculty of Transportation Engineering, Hu'nan Institute of Traffic Engineering, Hengyang 421000, China)

Abstract: With the development of smart grid, the power consumption of each scenario becomes more diversified, and accurate station load forecasting is the key to ensure that the relevant power sector to develop appropriate maintenance tasks, while providing an important reference for orderly power consumption and economic operation. In order to mine the characteristics of the station load to improve the accuracy of the station load forecasting, a station power load forecasting based on the kernel principal components analysis combined with local preservation projection for dimensionality reduction, K-means clustering algorithm (K-means), and bi-directional long short-term memory network (BiLSTM) is proposed. Firstly, the kernel local preservation projection (KLPP) is used to reduce the dimensionality of multi-featured load data in the station area to extract the main feature information. Secondly, the K-means clustering method is adopted to classify the data with similar features into their respective cluster sets. Finally, for each typical type after clustering, BiLSTM is trained in a targeted way, and the load of a low-voltage station area of a university in China is selected as an example to be compared and analysed with other classical forecasting methods. The proposed method is more suitable for the actual load direction and

effectively improves the prediction effect.

Key words: power load forecasting; dimensionality; K-means clustering algorithm; bi-directional long short-term memory networks; kernel local preservation projection

摘 要: 随着智能电网的发展,各场景的用电更加多元化,而准确的台区负荷预测是确保相关电力部门制定合适检修任务的关键,同时为有序用电、电网经济运行提供重要参考。为了挖掘台区负荷的特征以提高台区负荷预测的精度,提出了一种基于核主元分析与局部保持投影降维、K 均值聚类算法(K-means)以及双向长短期记忆网络(BiLSTM)的台区电力负荷预测方法。首先利用核局部保持投影(KLPP)对台区多特征负荷数据进行降维以提取主要特征信息;然后采取 K-means 聚类算法将相似特征的数据归类成各自的簇集;最后针对聚类后的各典型类型,有针对性地训练 BiLSTM,并选取中国某高校低压台区负荷作为算例与其他经典预测方法进行对比分析,结果表明所提方法更拟合实际负荷走向,有效提升了预测效果。

关键词: 电力负荷预测;降维;K 均值聚类算法;双向长短期记忆网络;核局部保持投影

0 引言

随着社会经济的发展我国用电需求增加以及“双碳”^[1]目标下“绿色竞争力”的增加,带来了建设以新能源为主体的新型电力系统的新需求,海

基金项目: 国家自然科学基金项目(61572416);湖南省自然科学基金项目(2022JJ50132)

National Natural Science Foundation of China (61572416);
Natural Science Foundation of Hu'nan Province, China (2022JJ50132)

量用户侧从低压台区入网^[2], 承接了庞大的电力数据量信息^[3]。因此对电力负荷信息的特征分析与预测提出了更高的要求。负荷预测可分为系统级与台区级。系统级负荷预测的目标是为各级调度部门制定调度计划提供依据; 台区级负荷预测的目标是指导检修任务、同时预报发电状况与电网动态负荷^[4]以及预报设备情况, 从而调整设备以减少运营成本。系统级涵盖的负荷种类多, 整体规律更显著; 台区负荷地域规模更小, 规律较弱。挖掘台区负荷数据的模态与规律能为决策人员提供支持, 因此台区电力负荷预测方法的研究受到了国内外学者的关注^[5]。

面对现在更为复杂的海量电力用户结构, 以传统回归分析法^[6]为代表的传统负荷预测方法难以达到理想的预测精度, 因此目前组合预测方法得到了更广泛的应用。文献[7]通过核主元分析 (Kernel Principal Components Analysis, KPCA) 将原始风电数据集降维后输入到预测模型中, 提高了计算效率, 但 KPCA 未能兼顾数据的局部特征信息。文献[8]基于小波变换对模型输入的台区负荷进行预处理, 然后采用深度神经网络来预测净荷, 通过仿真验证了所提预测模型的准确性。文献[9]通过主成分分析 (Principal Components Analysis, PCA) 提取影响配电台区的负荷因素, 然后采取模糊 C 均值进行特征划分后来预测台区负荷, 通过算例分析发现所提策略可有效提升预测结果的准确率, 但 PCA 处理非线性问题较差, 不能全面提取关键信息。

综上所述, 本文提出一种基于 KPCA 与局部保持投影 (Local Preservation Projection, LPP) 相结合的核局部保持投影 (Kernel LPP, KLPP) 降维方法, 首先降低样本的复杂度, 使得众多影响因素的数据低维化, 提取出内部特征; 其次通过 K-means 聚类算法将经 KLPP 降维后的数据进行聚类, 从而得到不同类簇的样本分量; 最后将聚类操作后的各分量进行针对性训练, 预测得到各分量然后聚合得到预测结果。

1 降维理论与分析

1.1 核局部保持投影 (KLPP)

KPCA 本质上是 PCA 降维的一种变形, 但 PCA 降维在多维数据降至低维空间时, 未能照顾

局部数据的特征形式, 而且设定数据是线性的, 与实际因素不符。而 LPP 在负荷样本降维过程中能很好地保持负荷样本局部流行的特点, 提高降维后用于负荷聚类的效果, 在保留关键信息的同时降低数据维度, 使得适用于预测模型网络训练的数据集综合信息能力更强, 提升预测效果, 因此结合 KPCA 与 LPP 的优点采取 KLPP 进行降维处理^[10]。

基于 KLPP 的电力负荷降维流程如下:

步骤 1: 输入高维特征数据集 $R_{m \times n}$, 经过分布中心化处理后得到数据集 $\bar{L}$ 。

步骤 2: 将中心化处理后的数据集 $\bar{L}$, 经由非线性特征函数 Φ 转换到多维核 Hilbert 空间中, 最后获得数据集在多维空间中的数据 $\Phi(X)$ 。

步骤 3: 选择混合核函数, 经计算得到多种影响因素间的内积。

步骤 4: 采用符合对应空间的 k 近邻度量值, 计算得到负荷点的相似度连接矩阵 M , 计算式如下:

$$M = \begin{cases} \sqrt{1 - e^{-\frac{d^2}{\theta}}}, & L_i = L_j \\ \sqrt{e^{-\frac{d^2}{\beta}} - \omega}, & L_i \neq L_j \end{cases} \quad (1)$$

式中: θ 为调节系数, 取决于负荷样本的聚集程度; ω ($0 \leq \omega \leq 1$) 为平衡系数, 用来平衡调整不同样本点的间距; L_i 和 L_j 分别为 i 节点和 j 节点的拉普拉斯矩阵; d 为邻近点的距离度量; β 为全局调节参数。

步骤 5: 根据式 (2) 计算照顾全局特征的 KPCA 全局目标函数; 根据式 (3) 计算照顾局部流行特征的 LPP 局部目标函数。

$$\begin{aligned} F_G &= P_F^T C_F P_F = \\ P_F^T &\left[\frac{1}{n} \sum_{j=1}^n \Phi(x_j) \Phi^T(x_j) \right] P_F = \\ A^T \Phi^T(X) &\left[\frac{1}{n} \sum_{j=1}^n \Phi(x_j) \Phi^T(x_j) \right] \Phi(X) A = \\ &\frac{1}{n} (A^T K K A) \end{aligned} \quad (2)$$

$$\begin{aligned} F_L(P_\phi) &= \sum_i \left(\sum_j \| P_\phi^T \Phi(x_i^c) - P_\phi^T \Phi(x_j^c) \|^2 M_L \right) = \\ &2\text{trace}[A^T K (D_L - M_L) K A] = \\ &2\text{trace}(A^T K L_L K A) \end{aligned} \quad (3)$$

式中: F_G 为 KPCA 的目标函数; F_L 为 LPP 的目标函数; Φ 为非线性特征函数矩阵; P_F 为第 F 个投影矩阵; X 为样本元素; C_F 为样本协方差; x_i 和 x_j 为数据样本点; x_i^c 和 x_j^c 分别为 x_i 的第 i 个和 j 个邻近的点; A 为系数矩阵; P_ϕ 为投影矩阵; trace 表示矩阵的迹; K 为核函数; D_L 为对角矩阵; M_L 为连接矩阵; L_L 为拉普拉斯矩阵。

步骤 6: 根据式(4)得到负荷降维过程中提炼主元的最佳函数。

$$F(a) = \underset{a}{\operatorname{argmax}} \left[\frac{\alpha}{n} (A^T K K A) - 2\beta \operatorname{trace}(A^T K L_L K A) \right] \quad (4)$$

s. t. $A^T K A = I$

式中: a 为目标函数线性组合系数矩阵; α 为局部调节参数; I 为单位矩阵。

步骤 7: 基于式(4)得到的最终映射负荷数据集, 使用特征分解求出相应的特征值 $\lambda_1, \lambda_2, \dots, \lambda_n$ 及特征值对应的特征向量 $\zeta_1, \zeta_2, \dots, \zeta_n$; 选择合适的降维主元计算出系数矩阵 A 及投影矩阵 P_ϕ 。

步骤 8: 输出降维结果。

1.2 KLPP 降维可视化对比分析

选取中国某高校台区 2018 年 8~10 月的用

电负荷数据, 包括气压、温度、云量、风速及湿度等五类负荷变化影响因素^[11]。影响因子的采样步长为 60 min, 一天 24 h, 每天采样样本为 120 个数据, 各影响因素的相关热力图如图 1 所示。由图 1 可知各影响因素与负荷呈现非线性关系从而证明后续采取非线性降维的合理性。为了验证所提 KLPP 降维方法的有效性, 选取 KPCA、LPP 以及 KLPP 等不同的降维方法对原始气象数据进行降维处理, 结果如图 2 所示。

由图 2(a) 与图 2(b) 可知, 使用 KPCA 降维法降到 2 维后不同类别有所区分, 但第 3 类特征

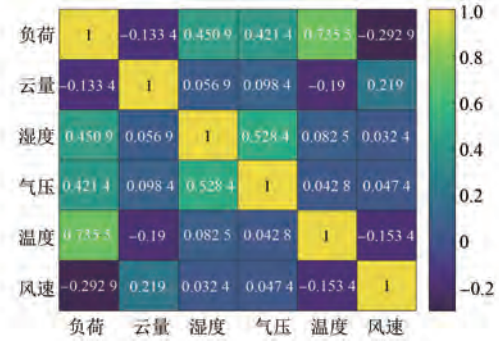


图 1 影响因素相关热力图

Fig. 1 Heat map related to impact factors

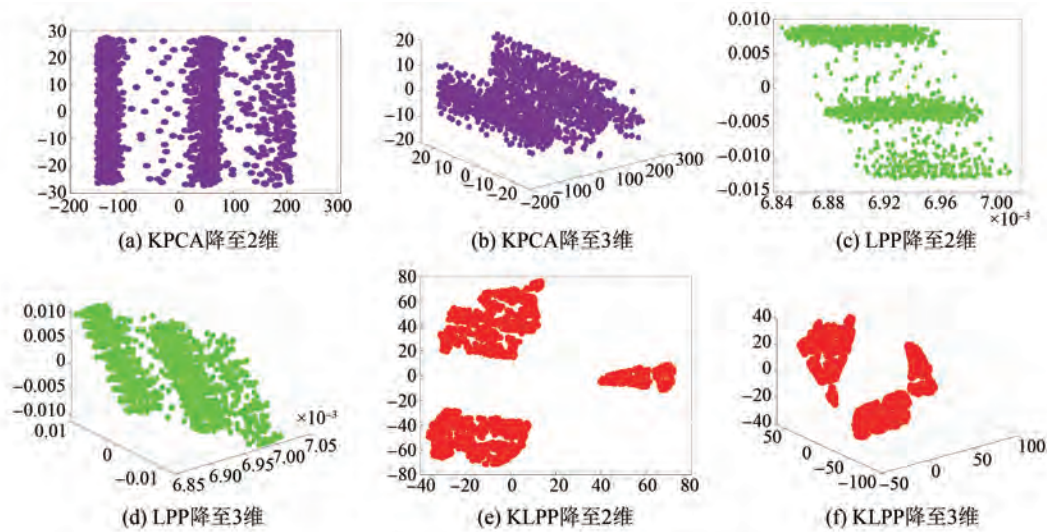


图 2 不同降维对比图

Fig. 2 Comparison chat of different downscaling

数据过于分散, 且降到 3 维后数据分散区别度不够; 由图 2(c) 与图 2(d) 可知, 相较于 KPCA, LPP 降维处理后 2 维中第 3 类特征数据聚集度有了一定的提升, 且降到 3 维后区分度有所提升; 由图 2

(e) 与图 2(f) 可知, 经过 KLPP 降维后不同类别数据样本边界距离清晰, 容易区分。其主要原因是 KPCA 降维仅考虑以整体分布方差为目标, 而 LPP 降维法则聚焦于保持数据局部流形结构不

变。因此 LPP 比 KPCA 降维结果的类内聚集性更好,但二者均不能兼顾样本整体方差与局部流形结构;而 KLPP 兼具 KPCA 和 LPP 的优点,因此数据降维效果最好,能挖掘出数据的隐藏特性^[12]。

2 基于 K-means 的聚类负荷方法划分

在针对台区电力负荷预测的研究中,研究者发现将相关特征的因素聚集起来,并进行分类,有针对性地对相关模型进行训练能提高负荷预测精度。校园台区负荷预测的研究通常包括温度、湿度等气象要素。本文选取重要程度不同的气象要素与负荷曲线,通过聚类分析划分出隶属度不同的类别,进行分类训练从而得到最优的网络超参数。本文采取 K-means^[13] 聚类算法,将降维后的原始气象数据通过加权欧式距离对负荷进行聚类。数据集构造为 $X = \{x_1, x_2, \cdots, x_n \mid x_i \in R^z\}$, 其中 z 为日特性因子的维度,建立的目标函数为

$$J(U, V, \omega, X) = \sum_{i=1}^k \sum_{j=1}^n u_{ij}^m d_{ij}^2 \tag{5}$$

式中: U 为一个 $k \times n$ 的矩阵; V 为聚合中心矩阵; ω 为特征权重矢量; m 是模糊加权指数; u_{ij} 为第 j 个数据归于第 i 类的模糊集合; d_{ij} 为不同样本聚集向量之间的欧式距离。

K-means 聚类算法在考虑未知标签的情况下,分离出不同的样本类型。因此需要提前给定簇心的 K 个样本,剩余样本则考虑与各簇心的距离,划分至最相同的簇心所属簇集。由于 K 值需要预置给定,而一般的判定 K 值的方法为手肘法,其运行效率低下,难以适应数据量庞大的工作。本文采用内部评估指标轮廓系数法^[14] 对簇内的凝聚度与分离度进行衡量评价,当轮廓系数越逼近 1 则表明簇内数据的聚集结果愈佳。轮廓系数表达式如下:

$$s = \frac{b - a}{\max(a, b)} \tag{6}$$

式中: a 为簇内凝聚度; b 为簇间分离度。

图 3~图 5 为 K 取不同值时的聚类结果。通过轮廓系数比较得到当 $K=3$ 时轮廓系数最大,因此本文选取 $K=3$,不同 K 值的轮廓系数如表 1 所示。

表 1 不同 K 值轮廓系数

Tab. 1 Contour coefficients for different K values

K 值	轮廓系数	K 值	轮廓系数
2	0.643 1	6	0.483 6
3	0.687 3	7	0.462 1
4	0.496 2	8	0.463 9
5	0.457 9	9	0.459 3

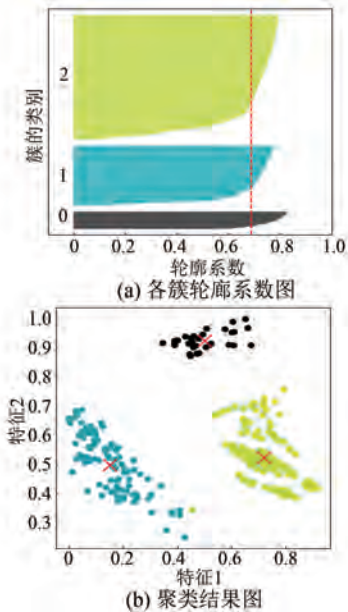


图 3 $K=3$ 对应的聚类结果

Fig. 3 Clustering results corresponding to $K=3$

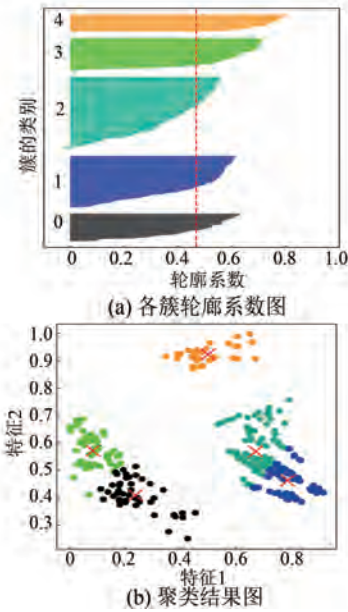
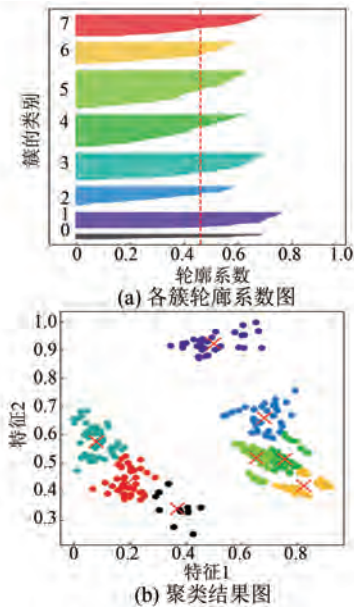


图 4 $K=5$ 对应的聚类结果

Fig. 4 Clustering results corresponding to $K=5$

通过 KLPP 对原始样本进行降维,将降维后

图5 $K=8$ 对应的聚类结果Fig. 5 Clustering results corresponding to $K=8$

的数据依照各样本之间的欧式距离划分所属类目,如式(7)所示:

$$d_i = \sqrt{\sum_{k=1}^3 [x_m(k) - x_i(k)]^2} \quad (7)$$

式中: $x_m(k)$ 为待预测对象降维后的数据; $x_i(k)$ 为历史训练样本数据。

以8月1日至18日为例,根据式(7)来划分样本所属类簇。在整个十八天样本当中有十一天归为第一类簇,四天归为第二类簇,三天归为第三类簇,如表2所示。

表2 不同历史负荷样本聚类结果($K=3$)Tab. 2 Results of clustering different historical load samples ($K=3$)

样本日	类别	样本日	类别
8月1日	1	8月10日	1
8月2日	3	8月11日	1
8月3日	1	8月12日	1
8月4日	1	8月13日	2
8月5日	3	8月14日	1
8月6日	2	8月15日	1
8月7日	1	8月16日	1
8月8日	1	8月17日	3
8月9日	2	8月18日	2

3 双向长短时记忆网络

双向长短时记忆 (Bi-Directional Long Short-Term Memory, BiLSTM) [15] 网络由一对方向相反

的长短时记忆 (Long Short-Term Memory, LSTM) [16] 层神经网络构成,因此具有双向学习的特性。这种特性使神经网络具有起点与终点同时递归训练的能力。前后 LSTM 构造的 BiLSTM 神经网络结构如图6所示。

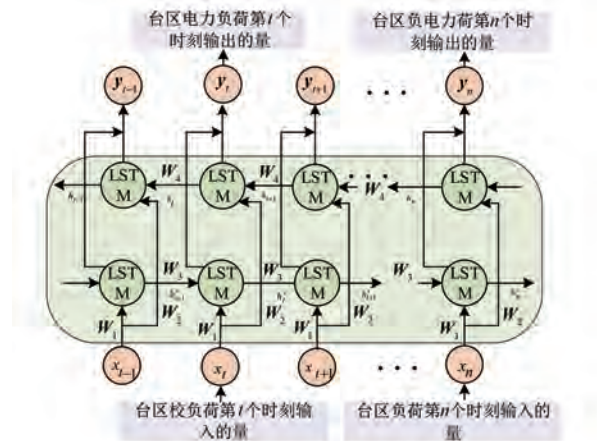


图6 BiLSTM 结构图

Fig. 6 BiLSTM structure diagram

由图6可知,前向 LSTM 层输出 $\vec{h}_t$ 与后向 LSTM 层输出 $\overleftarrow{h}_t$ 连接起来,产生 BiLSTM 的 y_t ,其更新方程为

$$\vec{h}_t = \text{LSTM}(x_t, \vec{h}_{t-1}) \quad (8)$$

$$\overleftarrow{h}_t = \text{LSTM}(x_t, \overleftarrow{h}_{t+1}) \quad (9)$$

$$y_t = W_{hy} \vec{h}_t + W_{hy} \overleftarrow{h}_t + b_y \quad (10)$$

式中: $\vec{h}_{t-1}$ 和 $\overleftarrow{h}_{t+1}$ 分别为神经网络上一时刻和下一时刻的输出值; W_{hy} 和 W_{hy} 为输出层连接权重矩阵; b_y 为输出层偏置。

由于台区电力负荷数据带有关联信息的时间特征,前后时间的负荷包含的信息具有联系性, LSTM 只能依赖前向数据,无法学习到后向信息。为了使神经网络充分学习到样本信息,捕获到上下样本间的信息,因此选择 BiLSTM 构建合适的台区电力负荷预测模型。模型中,隐含层个数 h 设置为3;学习率 ζ 设置为0.001;最大迭代次数设置为100; t 时刻台区电力负荷输入和输出分别为 x_t 和 y_t 。

4 台区短期电力负荷预测试验分析

4.1 台区短期电力负荷预测方法

首先通过 K-means 算法将降维后的台区电力

负荷数据集进行聚类处理以得到不同的类别;然后将不同的类别输入到预测模型中进行试验,KLPP-K-means-BiLSTM 台区电力负荷预测方法的流程如图 7 所示。

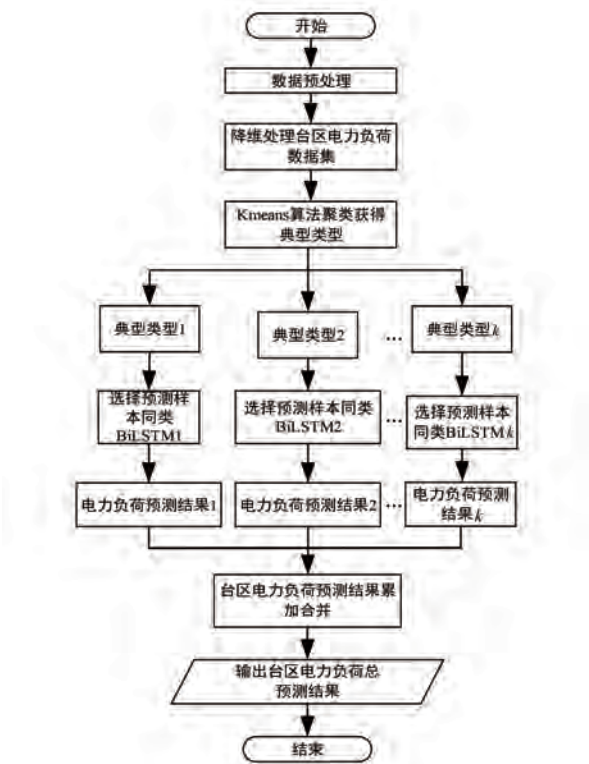


图 7 KLPP-K-means-BiLSTM 台区电力负荷预测方法流程图

Fig. 7 Flowchart of KLPP-K-means-BiLSTM power load forecasting method for station area

4.2 试验验证

为了验证本文所提 KLPP-K-means-BiLSTM 台区电力负荷预测方法的有效性,选择中国某高校 2018 年 8 月的台区电力负荷与气象数据作为样本来源,将最后七天作为测试样本,其他作为训练样本,台区电力负荷数据采样间隔为 1 h。降维对影响因素的贡献率如图 8 所示。从图 8 可以看出主要组成部分的贡献率和累计贡献率主要集中在前三类。因此选取前三类成分可以涵盖影响成分所有信息。同时为了验证台区电力负荷预测模型的有效性,选择支持向量机 (Support Vector Machine, SVM)、径向基函数 (Radial Basis Function, RBF) 神经网络以及 BiLSTM 不同预测方法进行对比,为减少随机因素对试验的干扰,每组试验运行 10 次,不同预测方法得到的预测结果对

比如图 9 所示。不同预测方法在试验 10 次后的平均误差指标的评价结果如表 3 所示,其中指标包括均方根误差 (Root Mean Square Error, RMSE)、平均绝对百分比误差 (Mean Absolute Percent Error, MAPE)^[17-18],RMSE 和 MAPE 的表达式分别为

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n [Y'(t) - Y(t)]^2} \quad (11)$$

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Y'(t) - Y(t)}{Y(t)} \right| \times 100\% \quad (12)$$

式中: n 为预测负荷的数量; $Y'(t)$ 为预测负荷; $Y(t)$ 为实际负荷。

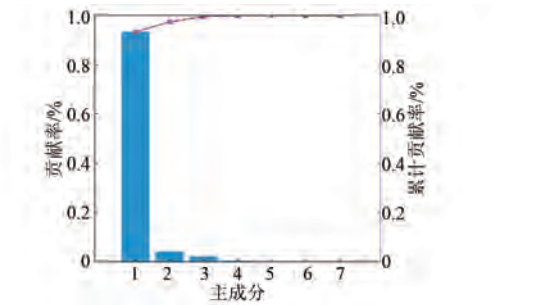


图 8 主成分贡献率及累计贡献率
Fig. 8 Principal components contribution and cumulative contribution

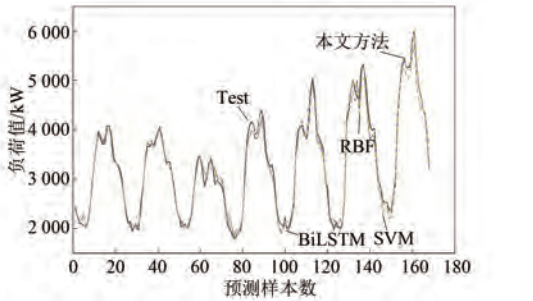


图 9 不同预测方法对比结果
Fig. 9 Comparison results of different prediction methods

表 3 不同方法的 RMSE、MAPE 对比
Tab. 3 Comparison of RMSE and MAPE for different methods

预测方法	RMSE	MAPE
SVM	152.613	3.541 0
RBF	130.632	3.150 1
BiLSTM	110.173	2.921 6
本文方法	101.212	2.360 9

从表 3 与图 9 中可以看出,本文所提方法的效果最好,MAPE 为 2.360 9, RMSE 为 101.212, 预

测误差比其他三种方法更小。验证了所提方法在改善台区电力负荷预测精度、降低预测误差上的有效性。

为了更直观地展示不同方法的预测效果将十次试验的 MAPE 与 RMSE 用箱型图表示,如图 10 和图 11 所示。可见,本文所提方法的箱型位置更低,表明误差更小;且箱型图上限与下限的距离更窄,表明误差更稳定。因此所提方法能进一步提升台区电力负荷预测方法的泛化能力。

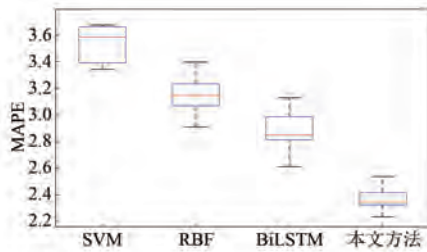


图 10 不同预测方法 MAPE 比较结果

Fig. 10 Comparison results of MAPE for different prediction methods

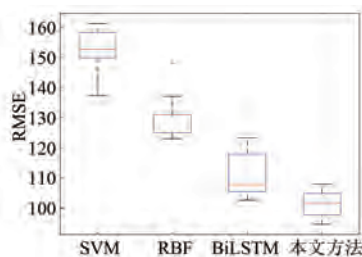


图 11 不同预测方法 MAPE 比较结果

Fig. 11 Comparison results of MAPE for different prediction methods

5 结语

为了提高台区电力负荷预测的精度,提出了一种 KLPP-K-means-BiLSTM 的台区短期电力负荷预测方法。得到如下结论:

(1) KLPP 降维方法兼顾了 KPCA 处理数据全局结构与 LPP 局部特性保持的优点;相比于 KPCA 和 LPP, KLPP 在降低非线性数据原始样本的复杂度上表现更优。

(2) 通过 K-means 算法将降维后的台区数据进行聚类处理以得到不同的类别;然后将不同的类别输入到预测模型中进行试验,通过 K-means 聚类算法将降维后的数据进行聚类,从而得到不同类别相似特性的样本分量,提高了模型的准确度。

(3) 与经典的 SVM、RBF 以及 BiLSTM 方法进行试验对比,验证了所提方法可提升台区电力负荷预测效果,减小误差,提高模型的泛化能力。

参考文献

- [1] 邓秋根,曹建华. 基于“双碳”战略的新型电力系统储能方案研究[J]. 电工技术, 2023, 10: 49-51.
DENG Q G, CAO J H. Research on new power system energy storage scheme based on double carbon strategy [J]. Transactions of China Electrotechnical Society, 2023, 10: 49-51.
- [2] 张勇军,羿应棋,李立涅,等. 双碳目标驱动的新型低压配电系统技术展望[J]. 电力系统自动化, 2022, 46(22): 1-12.
ZHANG Y J, YI Y Q, LI L Y, et al. Prospect of new low-voltage distribution system technology driven by carbon emission peak and carbon neutrality targets [J]. Automation of Electric Power Systems, 2022, 46(22): 1-12.
- [3] 葛磊蛟,范延赫,来金钢,等. 面向低碳经济的人工智能赋能微电网优化运行技术[J]. 高电压技术, 2023, 49(6): 2219-2238.
GE L J, FAN Y H, LAI J G, et al. Artificial intelligence enabled microgrid optimization technology for low carbon economy [J]. High Voltage Technology, 2023, 49(6): 2219-2238.
- [4] 魏国旺. 基于智能数据处理的电网信息实时交互系统设计[J]. 电子设计工程, 2023, 31(19): 94-98+103.
WEI G W. Design of a real-time interactive system for power grid information based on intelligent data processing [J]. Electronic Design Engineering, 2023, 31(19): 94-98+103.
- [5] 王丹,杨俊涛,曾燕,等. 考虑源荷不确定性的配电网风光储综合规划[J]. 电力需求侧管理, 2023, 25(5): 65-72.
WANG D, YANG J T, ZENG Y, et al. Comprehensive planning of wind-PV-storage in distribution network considering uncertainty of source and load [J]. Power Demand Side Management, 2023, 25(5): 65-72.
- [6] 郑乐,徐青山,冯小峰. 基于层次聚类算法与 ISA-LSSVM 的短期负荷预测研究[J]. 电力需求侧管理, 2022, 24(5): 51-57.
ZHENG L, XU Q S, FENG X F. Research on short term load forecasting based on hierarchical clustering

- algorithm and ISA-LSSVM [J]. Power Demand Side Management, 2022, 24(5): 51-57.
- [7] 徐艳, 周建勋, 金鑫, 等. 基于 KPCA-K-means-GRU 的短期风电功率预测研究[J]. 电机与控制应用, 2023, 50(2): 49-55.
- XU Y, ZHOU J X, JIN X, et al. Study on short-term wind power prediction based on KPCA-K-means-GRU [J]. Electric Machines & Control Application, 2023, 50(2): 49-55.
- [8] ALIPOUR M, AGHAEI J, NOROUZI M, et al. A novel electrical net-load forecasting model based on deep neural networks and wavelet transform integration [J]. Energy, 2020, 205: 118106.
- [9] 习伟, 蔡田田, 张镇, 等. 融合多元影响因素的配电台区 BiLSTM 负荷预测方法[J]. 天津大学学报 (自然科学与工程技术版), 2023, 56(11): 1205-1216.
- XI W, CAI T T, ZHANG Z, et al. BiLSTM load forecasting method for transformer districts integrated with multiple influencing factors [J]. Journal of Tianjin University (Science and Technology), 2023, 56(11): 1205-1216.
- [10] 杜董生, 盛远杰, 赵环宇, 等. 基于 VMD 的质子交换膜燃料电池系统故障诊断[J]. 控制工程, 2023, 30(7): 1190-1197.
- DU D S, SHENG Y J, ZHAO H Y, et al. Fault diagnosis of proton exchange membrane fuel cell system based on VMD [J]. Control Engineering of China, 2023, 30(7): 1190-1197.
- [11] 刘可真, 梁玉平, 代莹皓, 等. 基于 CNN-SSA-GRU 模型的短期负荷预测方法[J]. 电机与控制应用, 2022, 49(8): 79-86.
- LIU K Z, LIANG Y P, DAI Y H, et al. Short-term load forecasting method based on CNN-SSA-GRU model [J]. Electric Machines & Control Application, 2022, 49(8): 79-86.
- [12] SUETTERLEIN J, MANZANO J, MARQUEZ A, et al. Extending an asynchronous runtime system for high throughput applications: A case study [J]. Journal of Parallel and Distributed Computing, 2022, 163: 214-231.
- [13] 陈臣鹏, 赵鑫, 毕贵红, 等. 基于 Kmeans-VMD-LSTM 的短期风速预测[J]. 电机与控制应用, 2021, 48(12): 85-93.
- CHEN C P, ZHAO X, BI G H, et al. Short-term wind speed prediction based on Kmeans-VMD-LSTM [J]. Electric Machines & Control Application, 2021, 48(12): 85-93.
- [14] REN Y H, CHEN B, LI A Z. Intelligent system for sports movement quantitative analysis [J]. Journal of Intelligent & Fuzzy Systems: Applications in Engineering and Technology, 2021, 40(2): 3065-3073.
- [15] 冯斌, 张又文, 唐昕, 等. 基于 BiLSTM-Attention 神经网络的电力设备缺陷文本挖掘[J]. 中国电机工程学报, 2020, 40(增刊1): 1-10.
- FENG B, ZHANG Y W, TANG X, et al. Power equipment defect record text mining based on BiLSTM-Attention neural network [J]. Proceedings of the CSEE, 2020, 40(S1): 1-10.
- [16] 韩雪源. 基于 LSTM 算法的高压交联电缆线路振荡波局部放电检测方法[J]. 电机与控制应用, 2022, 49(12): 41-46+80.
- HAN X Y. Oscillatory wave partial discharge detection method of high voltage cross-linked cable line based on LSTM algorithm [J]. Electric Machines & Control Application, 2022, 49(12): 41-46+80.
- [17] 李青, 张新燕, 马天娇, 等. 基于 SSA-CNN-BiGRU-Attention 的超短期风电功率预测模型[J]. 电机与控制应用, 2023, 50(5): 61-71.
- LI Q, ZHANG X Y, MA T J, et al. Ultra-short-term forecasting model of wind power based on SSA-CNN-BiGRU-Attention [J]. Electric Machines & Control Application, 2023, 50(5): 61-71.
- [18] 徐萌. 基于 MA-SVM 方法的短期光伏功率预测[J]. 电机与控制应用, 2022, 49(7): 104-111.
- XU M. Short-term photovoltaic power prediction based on MA-SVM method [J]. Electric Machines & Control Application, 2022, 49(7): 104-111.

收稿日期:2023-10-15

收到修改稿日期:2023-12-28

作者简介:

朱江(1996-),男,硕士,助理工程师,研究方向为电力系统及其自动化,657382286@qq.com;

*通信作者:朱江(1996-),男,硕士,助理工程师,研究方向为电力系统及其自动化,657382286@qq.com。

KLPP-K-means-BiLSTM Based Short-Term Power Load Forecasting for Station Areas

ZHU Jiang^{1*}, WANG Fan¹, CAO Chuntang¹, YI Lingzhi², ZOU Jiale³

(1. CRRC Electric Vehicle CO., Ltd., Zhuzhou 412000, China;

2. School of automation and electronic information, Xiangtan University, Xiangtan 411100, China;

3. Faculty of Transportation Engineering, Hu'nan Institute of Traffic Engineering, Hengyang 421000, China)

Key words: power load forecasting; dimensionality; K-means clustering algorithm; bidirectional long short-term memory networks; kernel local preservation projection

In this paper, a combined forecasting model of dimensionality reduction and clustering is constructed, and a station power load forecasting based on the kernel principal components analysis combined with local preservation projection for dimensionality reduction, K-means clustering algorithm, and bi-directional long short-term memory network (BiLSTM) is proposed.

Firstly, the kernel local preservation projection (KLPP) is used to reduce the dimensionality of multi-featured load data in the station area to extract the main feature information. Secondly, the K-means clustering method is adopted to classify the data with similar features into their respective cluster sets. Finally, for each typical type after clustering, BiLSTM is trained in a targeted way. The station area power load forecasting model framework is shown in Fig.1.

The proposed method is experimentally compared with the classical support vector machine (SVM), radial basis function (RBF), and BiLSTM methods, and the results are shown in Fig.2 and Tab.1. It is verified that the proposed method can enhance the effect of power load prediction in the station area, reduce the error and improve the generalisation performance of the model.

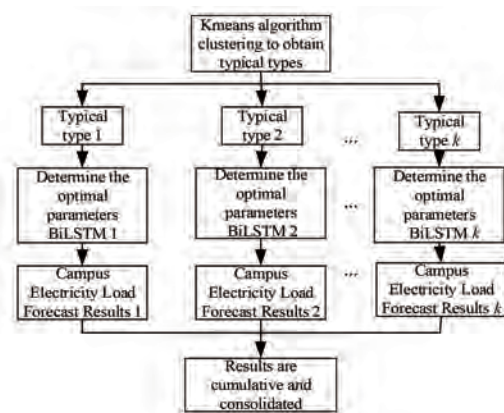


Fig. 1 KLPP-K-means-BiLSTM desktop power load forecasting model framework

Tab. 1 Comparison of RMSE and MAPE for different models

Model	RMSE	MAPE
SVM	152.613	3.5 410
RBF	130.632	3.1 501
BiLSTM	110.173	2.9 216
Proposed in this paper	101.212	2.3 609

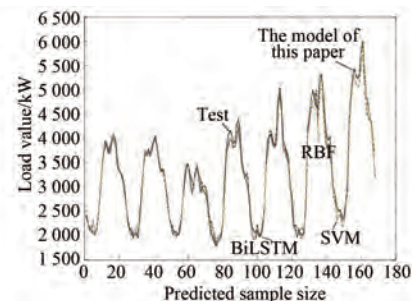


Fig. 2 Comparison results of different model prediction