

DOI: 10.19783/j.cnki.pspc.191097

# 基于特征选择和 Stacking 集成学习的配电网网损预测

邓威<sup>1</sup>, 郭钊秀<sup>2</sup>, 李勇<sup>2</sup>, 朱亮<sup>3</sup>, 刘定国<sup>3</sup>

(1. 国网湖南省电力有限公司电力科学研究院, 湖南 长沙 410007; 2. 湖南大学电气与信息工程学院, 湖南 长沙 410082; 3. 国网湖南省电力有限公司, 湖南 长沙 410004)

**摘要:** 针对配电网能量管理和节能降损的要求, 为了提高配电网网损分析与评估的有效性, 提出了一种基于特征选择和 Stacking 集成学习的配电网网损预测方法。首先基于特征选择的主要方法, 通过相关性分析法、最大信息系数法和基于树模型的特征选择法对特征进行综合分析, 得到各种特征对网损预测的重要性, 选择重要特征作为配电网网损预测模型的输入特征。在此基础上, 介绍 Stacking 集成学习原理, 考虑融合多种预测模型的优势特点, 建立 Stacking 集成学习配电网网损预测模型, 最后通过仿真验证得到网损预测结果。该仿真数据来源于湖南省 10kV 配电网某线路 44 个台区的真实数据, 网损预测结果表明该方法能够有效提升配电网网损预测的准确性和鲁棒性, 相比于单一预测模型具有更好的预测精度和泛化能力。

**关键词:** 特征选择; 模型融合; 集成学习; 网损预测

## Power losses prediction based on feature selection and Stacking integrated learning

DENG Wei<sup>1</sup>, GUO Yixiu<sup>2</sup>, LI Yong<sup>2</sup>, ZHU Liang<sup>3</sup>, LIU Dingguo<sup>3</sup>

(1. State Grid Hunan Power Company Limited Research Institute, Changsha 410007, China; 2. College of Electrical and Information Engineering, Hunan University, Changsha 410082, China; 3. State Grid Hunan Electric Power Company Limited, Changsha 410004, China)

**Abstract:** In order to improve the efficiency of distribution network losses analysis and evaluation on a distribution network, this paper proposes a distribution network loss prediction method based on feature selection and Stacking integrated learning. First, based on the main method of feature selection, through the correlation analysis method, the maximum information coefficient method and the tree-based model, the input features are comprehensively analyzed and the importance of various features for the network losses prediction is obtained. The important features are selected as the input of the distribution network losses prediction model. We then introduce Stacking integrated learning theory, which considers the advantages of combining multiple prediction models. We establish the Stacking integrated learning distribution network losses prediction model and finally obtain the results through simulation. The simulation data of the paper is derived from the real data of 44 stations in a line of a 10 kV distribution network in Hunan Province. The results show that the method can effectively improve the accuracy and robustness of distribution network losses prediction. Compared with a single prediction model, it has higher prediction accuracy and better generalization capabilities.

This work is supported by National Natural Science Foundation of China (No. 51822702), and Science and Technology Project of State Grid Hunan Electric Power Company (No. 5216A517000U), and the Excellent Innovation Youth Program of Changsha of China (No. KQ1802029).

**Key words:** feature selection; model fusion; integrated learning; losses prediction

## 0 引言

配电网网损是电力系统在配电中产生的电能损

**基金项目:** 国家自然科学基金项目资助(51822702); 国网湖南省电力有限公司科研项目资助(5216A517000U)“配电网损耗分析及降损增效辅助决策技术研究”; 长沙市杰出青年创新项目资助(KQ1802029)

耗, 是衡量电网运行质量、管理水平、经济效益的关键指标<sup>[1-2]</sup>。社会经济发展和用电负荷的增长导致配电网损耗也在增大。目前低压配电的网损已占整个电力网损耗的 40%左右<sup>[3]</sup>, 网损导致配电网运行的经济性下降, 但另一方面也说明配电网的节能降损具有很大的潜力和空间。因此, 精确地预测配电网网损, 可以为配电网出现量测故障的情况下提供数据支撑, 同时还可以反向检验量测系统数据的可

信度与准确性,及时发现异常量测值和故障量测装置;此外,预测得到的网损数据可以预估未来网损的变化情况,为配电网技改网改、配电网运行方式的调整和节能降损措施的实施提供辅指标。最后,配电网网损的预测结果还能协助电力系统进行规划设计和网络重构,对配电网节能降损、结构优化和电网经济效益提高具有重要的指导意义<sup>[4-5]</sup>。

近年来,智能电表和大数据采集存储系统的发展,为配电网网损预测提供了充足的数据支持。但是由于配电网网损数据的复杂性,现有的一些传统预测方法十分依赖于配电网网架结构,不能挖掘大量数据和配电网网损的内在规律。因此,一些统计学和人工智能算法逐步应用于配电网网损预测,利用这些算法的非线性拟合能力,能够较为精确地预测未来一段时间内配电网网损的变化情况,解决传统预测方法计算复杂度高,依赖配电网网架结构等问题。

目前一些研究人员已经提出了基于人工智能算法的配电网网损预测方法。在文献[6-7]中,研究人员分别采用人工神经网络和遗传算法来预测配电网网损,这些算法是建立在负荷预测和发电预测基础上进行的网损预测,是一种间接的配电网网损预测方法,在一定程度上可以较好地提升网损预测的精度,但同时也会存在误差积累问题。文献[8-9]中提出一种利用核心向量机和改进的神经网络算法来预测配电网网损的方法,提升了网损预测的精度。在文献[10]中,提出了一种基于灰色关联性分析的方法来筛选输入神经网络的电气特征,同时使用优化的 BP 神经网络来改善网损预测精度,该方法的预测输入特征与网损具有较强的相关性,能够提升配电网网损预测的准确性。

模型预测的方法可以对电力系统目前存在的问题提供指导<sup>[11-12]</sup>,但是目前上述方法都采用单一预测模型对配电网网损进行预测,由于配电网网损的复杂多变,单一预测模型自身也有一定的局限性,当网损随机变化时,可能导致预测效果不佳。此外,单一预测模型的精度提升有限,泛化性能较低。因此,为了克服单一模型的局限性,多模型融合的预测方法也应用在一些电力预测问题上。文献[13]采用对多种模型的预测结果赋予不同权重的方法来预测电力负荷,提升预测精度。文献[14]采用 Bagging 算法融合多组 SVM 模型进行负荷预测,其本质上是采用重采样和简单平均法来降低预测结果和真实值之间的方差,从而提高模型泛化能力。但是这些方法主要是通过线性化的方法来组合不同模型,所以并不能将各个模型的差异性和优势相融合。

针对配电网网损预测研究现状,本文提出了一种基于特征选择和 Stacking 集成学习的配电网网损预测方法。首先,对预测模型的输入特征进行选择,通过相关性分析、最大信息系数和基于树模型的特征选择方法,综合分析各个特征对配电网网损预测的重要性。分析结果表明历史网损、历史负载有功功率、小时、季度和节假日等特征具有较高的重要性,因此将这些特征作为网损预测模型的输入特征。在此基础上,构建了 K 最近邻预测模型(KNN)、支持向量机预测模型(SVM)、随机森林预测模型(RF)、梯度提升决策树预测模型(GBDT)和多层神经网络预测模型(MLP)分别对配电网网损进行预测。然后,结合各个单一模型建立 Stacking 集成学习网损预测模型,输入重要性分析所选择的特征数据,得到配电网网损预测的最终结果。对比分析单一模型和 Stacking 集成学习模型预测结果的准确性和泛化性,表明本文方法在网损预测方面的有效性和实用性。

## 1 输入特征重要性分析

### 1.1 特征选择主要方法

影响 10 kV 配电网网损的特征量复杂多样,需要对预测模型的输入特征进行筛选。特征选择可以去除无关特征,减轻输入数据维数过大的问题,同时降低预测模型的训练难度,提高预测精度。目前,特性选择主要是依据专家经验进行选择,没有统一的选择方法和标准。本文考虑特征选择的主要方法,结合采用相关系数法、最大信息系数法和基于树模型的特征选择方法来综合分析各特征对网损的重要性,确定预测模型输入特征。

#### 1.1.1 相关性系数法

相关系数法是一种能反映特征和目标变量之间相关性的统计方法<sup>[15]</sup>。常用的相关系数有皮尔森(Pearson)相关系数和斯皮尔曼(Spearman)相关系数。

Pearson 相关系数能够衡量变量之间的线性相关性,计算公式如式(1)所示。

$$\rho(x, y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

其中:  $n$  为样本数量;  $x$  为特征;  $y$  为目标变量;  $\bar{x}$  和  $\bar{y}$  分别为样本  $x$  和样本  $y$  的均值。Pearson 相关系数的取值范围为-1~1,相关系数的绝对值越大表示变量间的相关性越强,绝对值越小表示变量间的相关性越弱。但 Pearson 相关系数只对线性特征敏感,对于非线性对应关系的变量,无法采用 Pearson 相

关系数进行相关性分析。

Spearman 相关系数可以判断非线性变量间的相关性,它利用秩次的差值来评价两个变量间的非线性相关性。计算公式如式(2)所示,其中  $n$  为样本数量,首先对两个变量  $x_i$  和  $y_i$  的数据按照大小分别进行排序,然后记下排序以后的位置(即为秩次  $R(x_i)$  和  $R(y_i)$ )进行计算。Spearman 相关系数的取值范围为  $-1 \sim 1$ , 相关系数的绝对值越大表示变量间的相关性越强,绝对值越小表示变量间的相关性越弱。

$$\rho(x, y) = 1 - \frac{6 \sum_{i=1}^n |R(x_i) - R(y_i)|^2}{n(n^2 - 1)} \quad (2)$$

### 1.1.2 最大信息系数法

最大信息系数法(MIC)是通过计算联合概率来衡量变量间的相关性,该方法计算复杂度低,鲁棒性高,在负荷预测的特征选择方面已经获得较好的应用<sup>[16]</sup>。

最大信息系数的具体计算过程如式(3)所示。在二维空间  $x$  和  $y$  方向上分别划分  $a$  和  $b$  数量的网格,计算变量  $x$  和变量  $y$  在每个网格中的分布概率,最后通过标准化获得最大信息系数  $MIC$ 。其中,  $n$  为特征样本数,  $MIC$  的取值范围为 0 到 1, 最大信息系数越大,变量相关性就越强,最大信息系数越小,则相关性越弱。

$$MIC(x, y) = \max_{a*b < n^{0.6}} \frac{\int p(x, y) \log_2 \frac{p(x, y)}{p(x)p(y)} dx dy}{\log_2 \min(a, b)} \quad (3)$$

### 1.1.3 基于树模型的特征选择法

基于树模型的特征选择法是一种信息增益算法<sup>[17]</sup>,其原理是某特征包含的同类叶子节点数越多,则说明该特征在训练时的作用越显著,因此树模型能够通过学习训练来计算特征的重要程度并计算出特征贡献度指标。其中,特征贡献度指标是指树模型中每一棵树由特征形成分支节点的 Gini 指数下降程度之和。

常用于特征选择的树模型是 GBDT,它能挖掘输入数据的特征重要性,输出对应特征的贡献度指标。贡献度指标的取值范围是 0 到 1,特征贡献度指标越大,说明该特征在模型训练预测时越重要。

## 1.2 特征的重要性分析

本文选取湖南省 10 kV 配电网某线路的真实数据进行特征重要性分析,参与分析的特征为历史网

损、历史负载有功功率、历史负载无功功率和时间特征。其中,历史负载有功功率、历史负载无功功率和时间特征都为真实数据,但由于该线路电量量测数据不足,无法直接通过量测数据得到配电网网损数据。因此,本文采用电力系统商业仿真分析软件 Powerfactory,根据真实的配电网拓扑结构搭建此条线路的仿真模型,如图 1 所示。其中,标号圆形为 10 kV 线路中的各个台区,每个台区内有 10/0.4 kV 的变压器以及低压侧的等效负荷。将真实的配电网的电压、电流和负荷功率数据导入仿真模型,通过 Python 编写程序对每 30 min 的一组数据进行潮流仿真计算,得到 10 kV 配电网的网损数据。

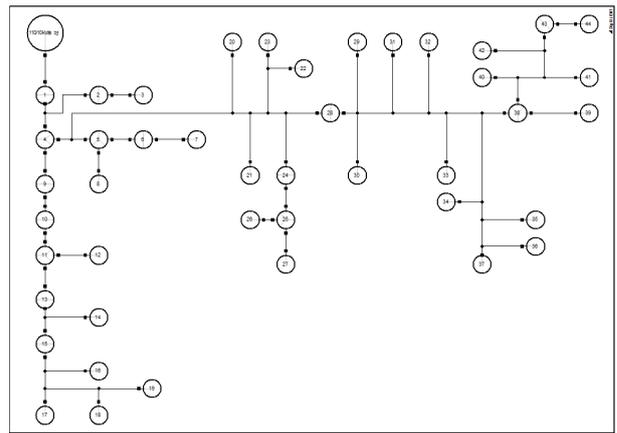


图 1 配电网线路仿真模型

Fig. 1 Line simulation model of distribution network

为了解决单一特征选择方法对某些特征数据不敏感的问题,因此采用相关系数、最大信息系数和基于树模型的特征选择三种方法,通过计算和训练得到特征与网损的相关系数、最大信息系数和特征贡献度指标,综合分析特征对网损预测的重要性。

特征分析的结果如图 2 所示,表 1 的编号和各种特征相对应。其中,1-7 号为前 1-7 天的历史网损,8-12 号为前 1-5 个小时的历史网损,13-14 号为前 1 天和前 1 小时的历史负载有功功率,15-16 号为前 1 天和前 1 小时的历史负载无功功率,17-21 号分别对应小时、星期、月份、季度和节假日特征。

从图 1 的重要性结果可知,在 Pearson 和 Spearman 相关系数法中,历史网损、历史负载有功功率的相关系数最大,说明这两个特征与网损的相关性最强。并且历史网损特征符合在实际情况中随时间推移历史网损与当前网损的相关系数依次下降、相关性依次减弱的变化规律;此外,小时、季度和节假日时间特征也具有较大的相关系数,说明

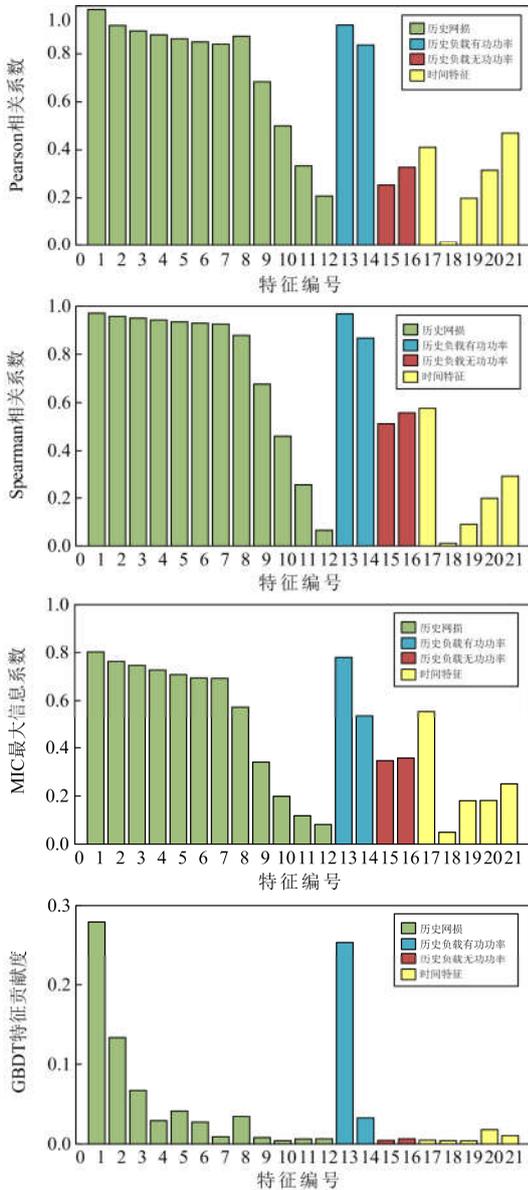


图 2 特征重要性结果

Fig. 2 Result of features importance

表 1 特征编号

Table 1 Number of features

特征	编号
历史网损(前 1—7 天)	1-7
历史网损(前 1—5 小时)	8-12
历史负载有功功率	13
历史负载无功功率	15-16
小时	17
星期	18
月份	19
季度	20
节假日	21

这些特征与网损也具有较强的相关性；历史负载无功功率的 Pearson 系数较小，Spearman 系数较大，说明历史负载无功功率与网损有一定的非线性相关性。在最大信息系数法中，历史网损、历史负载有功功率和小时的 MIC 最大信息系数最大，说明其与网损具有强相关性。在基于 GBDT 树模型的特征选择法中，历史网损和历史负载有功功率的特征贡献度指标最高，历史负载无功功率的特征贡献度指标很低，因此应该筛除历史负载无功功率特征。此外，由于树模型对离散型特征不敏感<sup>[18]</sup>，所以时间等离散特征的特征贡献度指标不高，因此需要通过相关性分析法和最大信息系数法对时间特征进行选择。

上述三种方法都表明历史网损和历史负载有功功率对网损预测的强重要性，小时、季度和节假日特征也与网损具有较强的关联性。因此，通过对输入特征的重要性分析，应选择历史网损、历史负载有功功率、小时、季度和节假日这些特征作为网损预测的输入特征。

## 2 基于 Stacking 集成学习的预测模型

### 2.1 基于 Stacking 集成学习预测原理

Stacking 集成学习预测方法如图 3 所示，初级学习器由各单一预测模型构成，首先利用原始数据集训练多个初级学习器。在训练过程中，为降低模型过拟合风险，一般采用  $K$  折交叉验证的方法对初级学习器进行训练。然后将初级学习器的预测结果构成新的数据集，再训练次级学习器，从而得到最终的预测结果，算法的具体步骤如下：

(1) 将原始数据集划分为两部分：原始训练集  $D$  和原始测试集  $T$ 。

(2) 对初级学习器进行  $K$  折交叉验证：将原始训练集  $D$  随机分成  $K$  等份  $(D_1, D_2, \dots, D_k)$ ，每个初级学习器将其中的 1 份作为  $K$  折测试集，剩下的  $K-1$  份作为  $K$  折训练集。使用  $K$  折训练集训练每个初级学习器，并对  $K$  折测试集进行预测，合并每个初级学习器的预测结果，作为次级学习器的训练集  $\tilde{D}$ 。

(3) 每个初级学习器对原始测试集  $T$  进行预测，将预测结果求平均作为次级学习器的验证集  $\tilde{T}$ 。

(4) 次级学习器从初级学习器中得到生成的新数据集：训练集  $\tilde{D}$  和验证集  $\tilde{T}$ ，然后进行学习训练，输出最终的预测结果。

Stacking 集成学习预测方法采用  $K$  折交叉验证来降低模型的过拟合风险，并且使用多种初级学习器的预测结果再进行次级训练，通过这种方法可以克服单一学习器的局限性，综合各种学习器的适用范围和优势特点，是一种提升预测结果准确性和泛化性的机器学习方法。

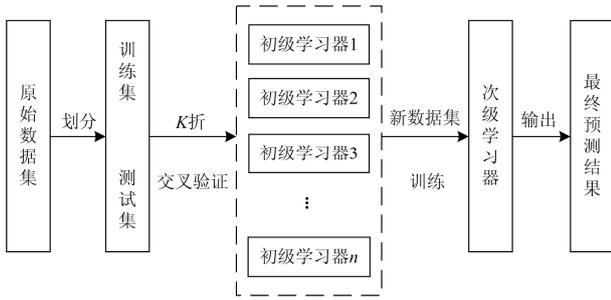


图3 Stacking 集成学习方法

Fig. 3 Method of Stacking integrated learning

## 2.2 预测模型评价指标

为了比较模型的预测效果, 本文采用平均绝对误差(MAPE)和均方根误差(RMSE)作为误差评价指标<sup>[19-21]</sup>, 具体计算如式(4)和式(5)所示。

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \tilde{y}_i|}{y_i} \times 100\% \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2} \quad (5)$$

其中:  $n$  为样本数量;  $y_i$  为真实值;  $\tilde{y}_i$  为预测值。

此外, 采用计算平均绝对误差的标准差 ( $\sigma MAPE$ ) 来说明预测误差的波动程度, 检验模型的稳定性和鲁棒性。计算如公式(6)所示。其中,  $n$  为样本数量,  $m_i$  为第  $i$  个预测样本的平均绝对误差,  $\bar{m}$  为  $n$  个样本的平局绝对误差的平均值。

$$\sigma MAPE = \sqrt{\frac{1}{n} \sum_{i=1}^n (m_i - \bar{m})^2} \quad (6)$$

## 3 算例验证与分析

本文采用湖南省 10 kV 配电网某线路 44 个台区 2017 年 1 月至 2018 年 7 月的真实数据, 建立预测模型对配电网网损进行预测。其中, 划分 2017 年 1 至 12 月共 365 天的数据作为原始训练集, 2018 年 1 至 7 月共 212 天的数据作为原始测试集。数据格式为每 30 分钟 1 个数据点, 一天共包含 48 个数据点。首先将数据进行预处理, 采用拉格朗日插值法填补缺失数据, 通过 Powerfactory 仿真得到网损数据, 按公式(7)将所有数据进行归一化处理, 归一化后的数据范围在 [0,1] 区间内。其中,  $X_{\max}$  和  $X_{\min}$  为数据集的最大值与最小值。

$$X^* = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (7)$$

### 3.1 单一网损预测模型

为对比 Stacking 集成学习预测模型的预测精度

和泛化性, 本文基于各学习器的预测能力, 选择预测性能较好的五种单一模型进行网损预测, 模型具体超参设置如表 2 所示。

在这五种单一模型中, KNN 理论完善, 训练高效, 它通过  $K$  个最近邻居样本输出的平均值来预测网损<sup>[22]</sup>; SVM 通过核方法优化损失函数和风险函数从而对网损进行预测, 它在小样本、非线性、高维度等回归预测方面具有广泛的应用<sup>[23-24]</sup>; RF 和 GBDT 都是基于树的预测模型, 它们分别采用 Bagging 和 Boosting 方式对网损进行预测, 并且模型不容易过拟合, 鲁棒性强<sup>[25-26]</sup>; MLP 是通过神经元的反向传播算法对网损进行预测<sup>[27]</sup>, 模型的拟合性能较好。

表 2 预测模型和超参设置

网损预测模型	超参设置
KNN	邻居数 $K=12$
SVM	径向核函数, 惩罚系数 1
RF	树的数量 300, 叶节点最小样本数 1, 节点划分最小样本数 2
GBDT	树的数量 180, 学习率 0.1, 树深度 7
MLP	2 层隐层, 神经元数量 50、20

本文通过表 2 的超参设置建立 5 个单一网损预测模型, 输入数据为特征重要性分析所选择的特征数据, 然后分别对网损进行训练和测试, 得到网损预测结果如表 3 所示。从表 3 中可知, 单一预测模型中 GBDT 的预测误差 MAPE 和 RMSE 最小分别为 4.13% 和 1.59, 说明 GBDT 的网损预测精度最高。此外, GBDT 模型的  $\sigma MAPE$  最小为 4.96, 说明其在单一预测模型中的稳定性能最好。

表 3 网损预测误差

网损预测模型	MAPE	RMSE	$\sigma MAPE$
KNN	4.45	1.63	5.18
SVM	4.86	1.83	5.80
RF	4.34	1.62	5.23
GBDT	4.13	1.59	4.96
MLP	4.85	1.69	5.34

### 3.2 Stacking 集成学习网损预测模型实例验证分析

为验证 Stacking 集成学习网损预测模型的性能, 本文集成 KNN、SVM、RF、GBDT、MLP 五种模型, 建立 Stacking 集成学习网损预测模型对网损进行预测。

本文建立的 Stacking 集成学习网损预测模型是一种集成多样化预测算法的综合预测模型, 能充分

利用各种算法从不同的角度和空间来预测网损。预测模型的具体结构如图 4 所示, 初级学习器由 KNN、SVM、RF、GBDT、MLP 预测模型构成, 次级学习器由预测精度较高, 稳定性较好的 GBDT 预测模型构成, 具体的训练流程如下:

1) 使用相关系数法、最大信息系数法和基于树的特征选择法对网损预测模型的输入特征进行重要性分析, 选择历史网损、历史负载有功功率、小时、季度和节假日作为网损预测模型的输入特征。

2) 对原始数据集按照 3.1 介绍的方法进行 5 折划分, 从图 3 可知,  $D_1$  到  $D_5$  中白色的数据集对应每个初级学习器的 5 折训练集, 灰色的数据集对应每个初级学习器的 5 折测试集。

3) 对初级学习器进行 5 折交叉验证, 训练 5 种初级学习器并输出预测结果, 组合生成新的数据集。

4) 采用新数据集对次级学习进行训练, 输出最终网损预测结果。

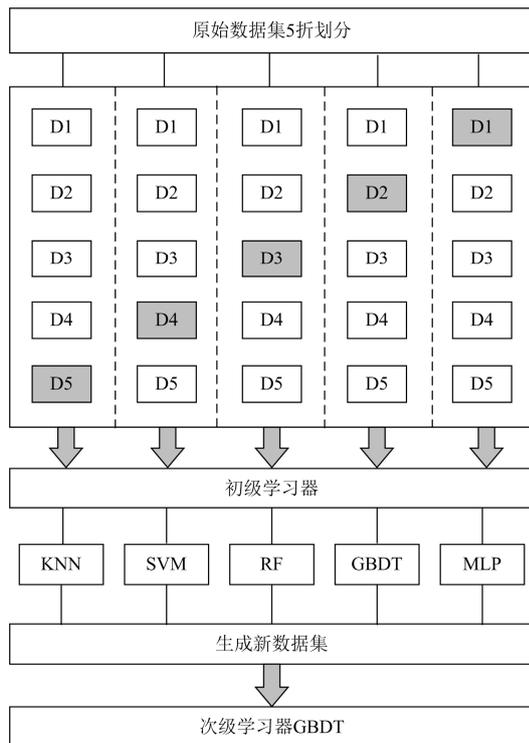


图 4 Stacking 集成学习网损预测模型结构

Fig. 4 Structure of Stacking integrated learning network losses prediction model

网损最终的预测结果如表 4 所示, 对比表 3 单一模型的预测误差可知: Stacking 集成学习预测模型预测结果的平均绝对误差和均方根误差为 3.78% 和 1.53, 相较单一模型中表现最好的 GBDT 还有所降低, 说明 Stacking 集成学习网损预测模型可以提

升网损预测的精度。同时, 集成学习预测模型的平均绝对误差标准差为 4.72, 说明预测误差波动最小, 验证了 Stacking 集成学习模型具有较强的稳定性。

表 4 网损预测误差

Table 4 Error of distribution network losses

网损预测模型	MAPE	RMSE	$\sigma$ MAPE
Stacking 集成学习预测模型	3.78	1.53	4.72

图 5 为单一网损预测模型和 Stacking 集成学习网损预测模型中三天的网损预测曲线, 其中红色曲线为真实网损, 蓝色曲线为 Stacking 集成学习模型的预测网损。从图中可知 Stacking 集成学习模型的预测结果精度更高, 更拟合真实网损曲线。此外, 第三天的网损曲线波动较大, 从图中可知单一预测模型对网损变化不敏感, 拟合效果较差, 然而 Stacking 预测模型仍能较好地拟合网损变化波动较大的曲线, 说明 Stacking 集成学习网损预测模型不仅能充分结合单一模型的预测优势, 而且还具有很强的泛化性能, 能较好地拟合不同波动情况下的网损曲线。在 2017 年 1 月至 2018 年 7 月的预测模型训练验证中, 该线路的两个专变台区出现过短时间的停运, 由于预测模型训练过程是对模型的参数进行计算优化, 优化结果使预测值能够拟合真实值, 所以在网络拓扑变化和负荷波动较大等复杂情况下, 本文预测方法仍能保证配电网网损的预测结果的精确与稳定, 说明 Stacking 集成学习网损预测具有较强的实用性。

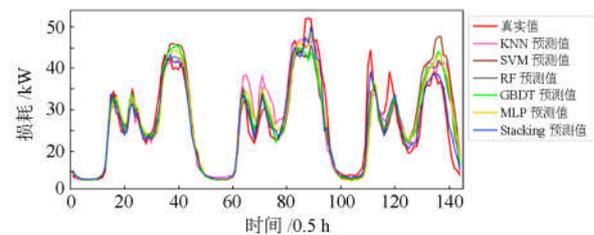


图 5 网损预测曲线

Fig. 5 Curve of distribution network losses prediction

## 4 结论

本文针对配电网网损预测的输入特征和预测效果展开研究, 提出了一种基于特征选择和 Stacking 集成学习的预测模型, 并通过实际的配电网数据进行实例验证。本文的主要结论如下:

1) 采用相关系数法、最大信息系数法和基于树的特征选择法综合确定网损预测模型的输入特征, 在进行网损预测时选择重要性强的特征作为输入能

够提高预测精度。

2) 结合单一模型的优势特点, 建立 Stacking 集成学习网损预测模型, 该模型具有很好的泛化性能, 能较好地挖掘网损数据规律, 提升网损预测精度和鲁棒性。

### 参考文献

- [1] 余健明, 蔡利敏, 杨文字, 等. 基于提高系统可靠性降低网损的配电网络重构[J]. 电工技术学报, 2004, 29(10): 70-73.  
YU Jianming, CAI Limin, YANG Wenyu, et al. Reconstruction of distribution network based on improving system reliability and reducing network loss[J]. Transactions of China Electrotechnical Society, 2004, 29(10): 70-73.
- [2] 张磊, 朱云鹤, 唐军, 等. 郴州 35 kV 及以上电网降低损耗工作现状分析及策略[J]. 湖南电力, 2017, 37(5): 23-28.  
ZHANG Lei, ZHU Yunhe, TANG Jun, et al. Analysis and strategy on the status of reducing loss of 35 kV and above power grids in Chenzhou[J]. Hunan Electric Power, 2017, 37(5): 23-28.
- [3] 李亚, 刘丽平, 李柏青, 等. 基于改进 K-Means 聚类 and BP 神经网络的台区线损率计算方法[J]. 中国电机工程学报, 2016, 36(17): 4543-4552.  
LI Ya, LIU Liping, LI Baiqing, et al. Calculation method of station line loss rate based on improved K-Means clustering and BP neural network[J]. Proceedings of the CSEE, 2016, 36(17): 4543-4552.
- [4] 余健明, 蔡利敏, 杨文字, 等. 基于提高系统可靠性降低网损的配电网络重构[J]. 电工技术学报, 2004, 29(10): 70-73.  
YU Jianming, CAI Limin, YANG Wenyu, et al. Distribution network reconstruction based on improving system reliability and reducing network loss[J]. Transactions of China Electrotechnical Society, 2004, 29(10): 70-73.
- [5] 张勇军, 翟伟芳, 林建熙, 等. 分布式发电并网的网损影响评价指标研究[J]. 电力系统保护与控制, 2011, 39(13): 134-137.  
ZHANG Yongjun, ZHAI Weifang, LIN Jianxi, et al. Research on the evaluation index of network loss impact of distributed generation grid[J]. Power System Protection and Control, 2011, 39(13): 134-137.
- [6] 陈章宝, 曾庆念, 徐耀, 等. 基于改进蚁群优化算法的配电网线损计算新方法[J]. 广东电力, 2012, 25(2): 72-76.  
CHEN Zhangbao, ZENG Qingnian, XU Yao, et al. A new method for calculating line loss of distribution network based on improved ant colony optimization algorithm[J]. Guangdong Electric Power, 2012, 25(2): 72-76.
- [7] 袁慧梅, 郭喜庆. 用遗传算法辅助设计的人工神经网络计算配电网线损[J]. 电网技术, 1998, 22(12): 19-21.  
YUAN Huimei, GUO Xiqing. Application of artificial neural network assisted by genetic algorithm to calculate line loss of distribution network[J]. Power System Technology, 1998, 22(12): 19-21.
- [8] 王齐. 基于支持向量机的线损计算及电能信息集成系统设计[D]. 长沙: 中南大学, 2011.  
WANG Qi. Line loss calculation based on support vector machine and design of power information integration system[D]. Changsha: Central South University, 2011.
- [9] 张勤, 周步祥, 林楠, 等. 基于灰色模型与神经网络组合的线损率预测[J]. 电力系统及其自动化学报, 2013, 25(5): 162-166.  
ZHANG Qin, ZHOU Buxiang, LIN Nan, et al. Prediction of line loss rate based on grey model and neural network combination[J]. Proceedings of the CSU-EPSA, 2013, 25(5): 162-166.
- [10] 张义涛, 王泽忠, 刘丽平, 等. 基于灰色关联分析和改进神经网络的 10 kV 配电网线损预测[J]. 电网技术, 2019, 43(4): 1404-1410.  
ZHANG Yitao, WANG Zezhong, LIU Liping, et al. Line loss prediction of 10 kV distribution network based on grey correlation analysis and improved neural network[J]. Power System Technology, 2019, 43(4): 1404-1410.
- [11] 冯兴田, 崔晓, 陶媛媛, 等. 一种基于改进预测电流控制的 Vienna 整流器设计[J]. 实验技术与管理, 2020, 37(6): 100-103, 184.  
FENG Xingtian, CUI Xiao, TAO Yuanyuan, et al. A design of Vienna rectifier based on improved predictive current control[J]. Experimental Technology and Management, 2020, 37(6): 100-103, 184.
- [12] 李文娟, 卮春颖, 张琦, 等. 永磁同步电机的扩展电压矢量模型预测控制[J]. 实验技术与管理, 2020, 37(2): 117-120, 131.  
LI Wenjuan, NI Chunying, ZHANG Qi, et al. Extended voltage vector model predictive control of permanent magnet synchronous motor[J]. Experiment Technology and Management, 2020, 37(2): 117-120, 131.
- [13] 谷云东, 张素杰, 冯君淑, 等. 大用户电力负荷的多模型模糊综合预测[J]. 电工技术学报, 2015, 30(23): 110-115.  
GU Yundong, ZHANG Sujie, FENG Junshu, et al. Multi-model fuzzy comprehensive prediction of large user power load[J]. Transactions of China Electrotechnical Society, 2015, 30(23): 110-115.
- [14] 张杰龙. 基于组合权重的中长期负荷预测研究及系统

- 实现[D]. 长沙: 湖南大学, 2016.
- ZHANG Jielong. Medium and long-term load forecasting based on combined weights and its system implementation[D]. Changsha: Hunan University, 2016.
- [15] 钟庆, 陈伟坤, 许中, 等. 设备故障统计数据与电能质量监测数据的关联分析[J]. 电力电容器与无功补偿, 2016, 37(6): 29-34, 45.
- ZHONG Qing, CHEN Weikun, XU Zhong, et al. Correlation analysis between equipment fault statistics and power quality monitoring data[J]. Power Capacitor and Reactive Power Compensation, 2016, 37(6): 29-34, 45.
- [16] CHU Fuchang, FAN Zhenping, GUO Baohua, et al. Variable selection based on maximum information coefficient for data modeling[C]// 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), March 25-26, 2017, Chongqing, China, 1714-1717.
- [17] 谢国荣, 郑宏, 林伟祈, 等. 基于改进随机森林算法的停电敏感用户分类[J]. 计算机系统应用, 2019, 28(3): 104-110.
- XIE Guorong, ZHENG Hong, LIN Weiqi, et al. Power outage sensitive customers classification based on improved random forest algorithm[J]. Computer Systems & Applications, 2019, 28(3): 104-110.
- [18] HAQ A U, ZHANG DEFU, PENG HE, et al. Combining multiple feature-ranking techniques and clustering of variables for feature selection[J]. IEEE Access, 2019, 27: 151482-151492.
- [19] 蒋峰, 王宗耀, 张鹏. 基于灰色-加权马尔可夫链的光伏发电量预测[J]. 电力系统保护与控制, 2019, 47(15): 55-60.
- JIANG Feng, WANG Zongyao, ZHANG Peng. Prediction of photovoltaic power generation based on gray-weighted Markov chain[J]. Power System Protection and Control, 2019, 47(15): 55-60.
- [20] 杨茂, 张罗宾. 基于数据驱动的超短期风电功率预测综述[J]. 电力系统保护与控制, 2019, 47(13): 171-186.
- YANG Mao, ZHANG Luobin. Overview of data-driven ultra-short-term wind power prediction[J]. Power System Protection and Control, 2019, 47(13): 171-186.
- [21] 肖逸, 谢俊, 刘若平, 等. 计及风电功率预测误差的备用容量计算新方法[J]. 电力系统保护与控制, 2019, 47(9): 72-79.
- XIAO Yi, XIE Jun, LIU Ruoping, et al. A new method for calculating the reserve capacity of wind power prediction error[J]. Power System Protection and Control, 2019, 47(9): 72-79.
- [22] 林芳, 林焱, 吕宪龙, 等. 基于均衡 KNN 算法的电力负荷短期并行预测[J]. 中国电力, 2018, 51(10): 88-94, 102.
- LIN Fang, LIN Yan, LÜ Xianlong, et al. Short-term parallel prediction of power load based on equilibrium KNN algorithm[J]. Electric Power, 2018, 51(10): 88-94, 102.
- [23] 牛东晓, 谷志红, 邢棉, 等. 基于数据挖掘的 SVM 短期负荷预测方法研究[J]. 中国电机工程学报, 2006, 26(18): 6-12.
- NIU Dongxiao, GU Zhihong, XING Mian, et al. Research on short-term load forecasting method of SVM based on data mining[J]. Proceedings of the CSEE, 2006, 26(18): 6-12.
- [24] 范昕炜. 支持向量机算法的研究及其应用[D]. 杭州: 浙江大学, 2003.
- FAN Xinwei. Research and application of support vector machine algorithm[D]. Hangzhou: Zhejiang University, 2003.
- [25] 王德文, 孙志伟. 电力用户侧大数据分析并行负荷预测[J]. 中国电机工程学报, 2015, 35(3): 527-537.
- WANG Dewen, SUN Zhiwei. Big data analysis and parallel load forecasting for power user side[J]. Proceedings of the CSEE, 2015, 35(3): 527-537.
- [26] 毕云帆, 撒奥洋, 张智晟, 等. 基于模糊 Bagging-GBDT 的短期负荷预测模型研究[J]. 电力系统及其自动化学报, 2019, 31(7): 51-56.
- BI Yunfan, YAN Aoyang, ZHANG Zhisheng, et al. Short-term load forecasting model based on fuzzy Bagging-GBDT[J]. Proceedings of the CSU-EPSA, 2019, 31(7): 51-56.
- [27] 冯凯. 基于大数据的居民用电行为分析与负荷预测[D]. 北京: 华北电力大学, 2017.
- FENG Kai. Analysis of residential electricity behavior and load forecasting based on big data[D]. Beijing: North China Electric Power University, 2017.

收稿日期: 2019-09-08; 修回日期: 2019-10-09

作者简介:

邓威(1983—), 男, 博士, 高工, 研究方向为智能配电网运行; E-mail: 13460103@qq.com

李勇(1982—), 男, 通信作者, 博士, 教授, 研究方向为电力系统优化与控制、电能质量分析与控制、电力电子系统与控制。E-mail: yongli@hnu.edu.cn

(编辑 葛艳娜)