

DOI: 10.19783/j.cnki.pspc.210580

一种 Stacking 集成结构的台风灾害下停电空间预测方法

侯慧¹, 陈希¹, 李敏², 朱凌², 黄勇³, 朱韶华¹

(1. 武汉理工大学自动化学院, 湖北 武汉 430070; 2. 广东电网有限责任公司, 广东 广州 510080;
3. 广东电网有限公司电力科学研究院, 广东 广州 510080)

摘要: 为提高电网公司防灾减灾能力, 考虑电网、气象、地理等因素, 从统计学习的角度提出一种台风灾害下停电空间预测方法。首先, 以 1 km×1 km 网格为单位收集数据, 并进行标准化、分类变量独热编码处理与筛选、构造特征等处理后作为模型输入数据。其次, 选取随机森林、梯度提升决策树、自适应提升、K 最近邻、支持向量机、极限树、决策树以及 XGBoost 等算法, 利用 Stacking 集成技术构造停电空间预测模型。最后, 以广东省某县为研究对象, 对模型的停电空间预测性能进行验证。在台风“彩虹”下的预测准确率为 0.777 6, 召回率为 0.914 0。结果验证了在台风灾害下该模型对停电空间预测的可行性与有效性。

关键词: 台风灾害; 统计学习; 机器学习算法; Stacking; 停电空间预测

A space prediction method for power outage in a typhoon disaster based on a Stacking integrated structure

HOU Hui¹, CHEN Xi¹, LI Min², ZHU Ling², HUANG Yong³, ZHU Shaohua¹

(1. School of Automation, Wuhan University of Technology, Wuhan 430070, China; 2. Guangdong Power Grid Co., Ltd., Guangzhou 510080, China; 3. Electric Power Research Institute of Guangdong Power Grid Co., Ltd., Guangzhou 510080, China)

Abstract: To improve disaster prevention and the mitigation ability of power grid companies, taking into account the power grid, meteorology and geography, a power outage spatial prediction method in the case of typhoon disasters is proposed from the perspective of statistical learning. First, data are collected in a 1 km×1 km grid, and standardized, classified with variable one-hot encoding and screening, and construction features are processed as input data for the model. Secondly, algorithms such as random forest, gradient boosting decision tree, adaptive boosting, K-nearest neighbor, support vector machine, extreme tree, decision tree, and XGBoost are selected to construct an outage space prediction model using Stacking integration technology. Finally, a city in Guangdong Province is taken as the research object to verify the performance of the model: the prediction accuracy rate for typhoon "Mujigae" is 0.777 6, and the recall rate is 0.914 0. The results verify the feasibility and effectiveness of the model in predicting power outage space in the case of typhoon disasters.

This work is supported by the University-Industry Collaborative Education Program of the Ministry of Education (No. 201902056044).

Key words: typhoon disaster; statistical learning; machine learning algorithm; Stacking; outage space prediction

0 引言

外部自然灾害一直是电网发生故障的主要原因^[1], 其中台风会造成电网发生大范围故障甚至大

停电事故, 严重威胁电网安全稳定运行, 同时造成重大经济损失^[2]。因此, 开展台风灾害下的停电预测研究, 对电网公司灾前风险量化评估、制定应急抗灾策略、提高电网公司灾害应急处置能力具有重要现实意义。

台风对电力系统影响的研究主要分为脆弱性建模和统计学习两方面^[3]。脆弱性建模是指建立设备在台风灾害下的脆弱性曲线模型, 基于模型研究台

基金项目: 教育部产学研合作协同育人项目资助(201902056044); 中国南方电网有限责任公司科技项目资助(GDKJXM20198441 (036100KK52190053))

风下的设备故障风险。在脆弱性建模方面,文献[4]结合脆弱性曲线与应力干涉模型对台风下的输电线路风险进行预警,但是计算效率不高且不适用于大数量设备下的风险预测。文献[5-6]考虑风荷载实时故障率及状态的动态交替特性,建立配网杆塔可靠性评估模型,但是仅考虑了风速大小而未考虑其他因素,如风向、微地形、老化等因素,显得模型过于理想。文献[7]从风速和风向联合的角度出发对电力断线倒塔概率预测模型进行建模,文献[8]利用蒙特卡洛法模拟风速大小和方向,建立输电线路风偏跳闸模型,但同时模型计算也更加复杂。文献[9]考虑地理高程信息,结合风荷载下的脆弱曲线提出输电线路台风下的预警模型。文献[10-11]考虑微地形因素,结合设备脆弱性曲线建立输电线路风险预警模型。虽然以上研究从不同角度加入了更多的因素进行考虑,但是同时也加大了模型计算的复杂性,因此一般仅适用于样本和变量较少的情况。另一方面,以上研究从建模理论出发,可能导致变量取值过于理想,实际应用受限。因此有必要寻找一种能够考虑多样本多变量且变量准确反映实际状态的方法,进行台风灾害下的停电风险评估研究。

人工智能算法的发展为学者研究极端灾害下的停电问题提供了另一个研究出发点,那就是从统计学习的角度对灾害下的电网风险进行评估研究,这种方法更灵活、更准确、更快速。在这方面,文献[12]利用逻辑回归建立气象因素对电力事故的影响模型,文献[13]基于模糊和相似度对输电线路可靠性参数进行预估,但是仅分析了气象方面的多因素,且未上升到对系统层面停电预测问题的研究。文献[14]利用负二项回归模型进行台风下的停电预测研究,[15-16]分别提出台风下停电预测的广义线性混合模型和广义加性模型,但是由于网格范围过大,预测的精度较低。文献[17]利用随机森林算法进行停电空间预测并利用权重法二次评估提高预测效果,文献[18]利用随机森林算法对停电区域和数量进行预测,虽然以上研究^[17-18]都实现了停电区域的预测,且预测精度为 $1\text{ km}\times 1\text{ km}$,但是上述模型仅从单一算法出发。对于停电事件而言,特征数据之间的空间结构关系非常复杂,单一算法仅从自身算法角度分析数据,限制了对特征数据之间复杂关系的充分挖掘,因此有必要采用集成多算法的方式进行数据挖掘。

文献[19]利用集成的方法对决策树、随机森林、增强梯度树回归三种模型集成并比较,通过对比表明,集成后的模型比单一模型在停电预测上具有更优异的效果。模型集成的效果在一定程度上取决于

集成的方法,Stacking 是机器学习三大集成方式之一,文献[20-22]利用 Stacking 的方法开展研究,结果表明,利用 Stacking 构建的集成模型能够在单一模型的基础上提升预测性能,因此利用 Stacking 集成的方法融合多算法,从统计学习的角度出发研究台风对电力系统的影响是一个值得探讨研究的方向。

因此,本文提出一种 Stacking 集成结构的台风灾害下停电空间预测方法:利用 Stacking 技术对多个算法进行融合,对 $1\text{ km}\times 1\text{ km}$ 网格划分后的区域进行停电空间预测,实现较高的预测精度。本文方法利用 Stacking 从不同算法的角度挖掘特征之间的关系,避免单一算法在特征关系表现上的单一关注,从而导致训练时出现偏置性,实现对模型预测性能的提升。

1 台风灾害下停电空间预测框架

本文建立的台风灾害下停电空间预测框架如图 1 所示。

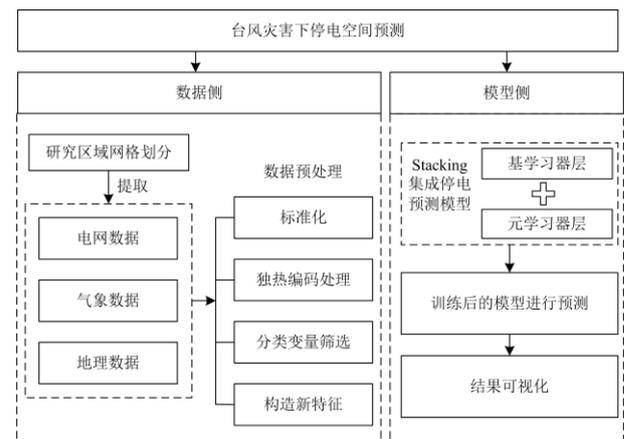


图 1 台风灾害下停电空间预测框架

Fig. 1 Spatial prediction framework for power outages under typhoon disasters

1) 考虑到实际数据收集可行性及精度要求,基于 $1\text{ km}\times 1\text{ km}$ 的精度对研究区域进行网格划分。

2) 以网格为单位收集样本数据,包括电网数据、气象数据、地理数据等,数据来源于电网公司。若存在所提供的数据精度与本文精度不相符的情况,利用 ArcGIS^[23]软件的插值功能进行转换提取。

3) 对收集的样本数据进行标准化、分类变量处理、构造特征变量等一系列预处理工作,作为模型的输入数据。

4) 基于 Stacking 技术构造包含基学习器层和元学习器层的两层 Stacking 集成模型,对模型进行训练,利用训练后的模型进行停电空间预测,最后利

用 ArcGIS 对结果进行可视化, 直观反映停电网格空间分布趋势。

2 样本数据收集与处理

2.1 网格单位数据收集

台风灾害下的区域停电事件是多方面因素作用的结果, 主要涉及电网、气象、地理三个方面^[17], 因此, 本文将研究区域进行 1 km×1 km 网格化划分后, 以网格为单位对这三个方面样本数据进行收集。

1) 电网数据: 台风导致倒塔断线是造成台风灾害下停电事件发生的主要原因之一, 因此选择与倒塔断线相关的杆塔数 V1、线路长度 V2 作为特征变量。对杆塔采取拉线措施能在一定程度上提升杆塔在台风下的抗风能力, 进而提升“保电能力”, 因此拉线加固杆塔数 V3 也作为特征变量。另外, 在收集数据时, 电网公司通过网格中的箱变数、台变数来确定网格内的用户数, 因此箱变数 V4、台变数 V5、以及用户数 V6 同样作为特征变量收集。

2) 气象数据: 台风灾害下发生倒塔断线的主要原因是构件在台风下受到的风荷载超过其设计风荷载, 因而发生倒塔断线导致停电。而风偏跳闸停电的主要原因为绝缘子串在大风条件下风偏角变小发生跳闸导致停电。这两类事件都与风速有关, 因此选取最大阵风风速 V7 作为特征变量。台风往往会造成降雨, 但雨荷载在实际中对杆塔的影响与风荷载对杆塔的影响相比很小, 且本文重点研究台风灾害, 对于台风导致的山洪等次生灾害不予考虑, 因此本文的气象数据中不考虑降雨量这一特征。

3) 地理数据: 风场中的不同位置, 其风速也不同, 经纬度信息可以确定某个点的位置信息, 因此将经度 V8、纬度 V9 作为特征变量。通过经纬度获取的风速仅仅是风速的观测值, 但是实际上具体到某根杆塔位置处, 其微地形是不同的, 而不同微地形对风速的影响不同, 如典型的“狭管效应”^[11], 因此本文考虑微地形对风速的影响, 选取海拔 V10、坡度 V11、坡向 V12 作为特征变量表征网格内的微地形差异。除地形因素外, 地表的差异同样会对风速造成影响, 文献[24]已证明地表差异对站点风速有较明显的影响, 因此在数据收集阶段, 本文选取地表类型 V13、下垫面类型 V14 作为特征变量表征不同网格之间的地表差异。

本文研究的是预测台风灾害下网格单位内是否发生因受灾而导致的用户停电, 是一个二分类问题, 因此输出是对网格是否停电的判断: 当网格内发生用户停电时, $U=1$; 网格内未发生用户停电时, $U=0$ 。综上所述, 本文的样本数据具体如表 1 所示。

表 1 原始样本数据类别与具体变量
Table 1 Original sample data categories and specific variables

类别	变量名称
电网数据	杆塔数 V1
	线路长度 V2
	拉线加固杆塔数 V3
	箱变数 V4
	台变数 V5
	用户数 V6
特征变量	气象数据
	最大阵风风速 V7
	经度 V8
	纬度 V9
	海拔 V10
	地理数据
	坡度 V11
坡向 V12	
	地表类型 V13
	下垫面类型 V14
因变量	网格是否发生停电 U(0 或 1)

2.2 数据预处理

在对数据进行训练之前, 有必要对数据进行预处理。由于收集的数据中, 不同特征变量归属于不同类型, 不同变量之间的量纲不同, 会对模型预测的可行性及可靠性产生影响, 为了消除特征变量之间的量纲影响, 需要对特征变量按式(1)进行标准化处理, 使得不同指标之间具有可比性。

$$V_{\text{nor}} = \frac{V - V_{\text{min}}}{V_{\text{max}} - V_{\text{min}}} \quad (1)$$

式中: V_{nor} 为变量 V 标准化后的值; V_{min} 、 V_{max} 分别为变量 V 的最小、最大值。

在所选取的特征变量中, 地表类型、下垫面类型属于分类变量, 分类变量仅代表类别或标签, 没有实际数值上的大小意义, 因此需要对分类变量进行处理。本文采用独热编码(One-hot encoding)^[25]的方法对其进行处理。另外, 在分类变量中, 地表类型有 25 种(0~24), 下垫面类型有 10 种(0~9), 对所有样本数据进行地表类型与下垫面类型的样本数量统计, 如图 2 所示。

如图 2 所示, 地表类型中, 类型 2(旱地)所占比例最大, 下垫面类型中, 类型 0(水稻土)所占比例最大。过多冗余特征加入会影响模型的预测性能, 因此, 为避免引入过多无用特征, 同时保留地表类型、下垫面类型这两个表征地表差异的特征, 选择“旱地”和“水稻土”作为地表类型和下垫面类型的特征变量。

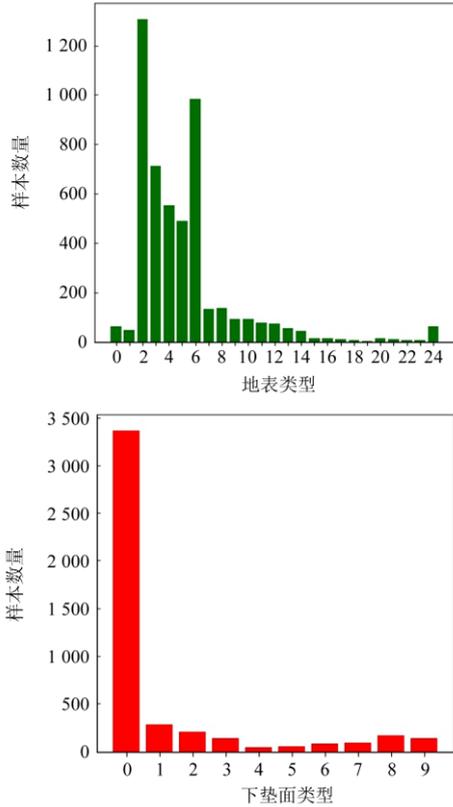


图2 地表类型与下垫面类型样本统计
Fig. 2 Sample statistics of surface type and underlying surface type

最后, 考虑到实际中拉线加固措施对杆塔在台风下的抗风能力提升具有增益效果, 会对杆塔倒塌引发断线从而导致用户停电产生影响, 因此, 利用“杆塔数 $V1$ ”与“拉线加固杆塔数 $V3$ ”构造“拉线加固影响 V_{new} ”这一变量来表征网格内拉线加固效应对停电的影响:

$$V_{new} = V1 - V3 \quad (2)$$

而考虑到“拉线加固杆塔数 $V3$ ”和“拉线加固影响 V_{new} ”都反映拉线加固效应, 且实际中拉线加固的线路数量较少, 变量取值上存在较多 0 值, 不利于信息挖掘, 而根据式(2)可知, “拉线加固影响 V_{new} ”不存在该问题, 因此以变量“拉线加固影响 V_{new} ”替代变量“拉线加固杆塔数 $V3$ ”。

3 Stacking 集成模型构建

3.1 算法选择

“集成学习”是一种建立在统计学习理论基础之上的多算法融合方法, 而 Stacking 集成是机器学习的三大集成方式之一^[26], 它是指组合多个算法的预测数据以生成新模型的集成技术。本文利用

Stacking 集成方法对多个算法进行融合, 构建包括基学习器层和元学习器层的两层 Stacking 集成模型。

在基学习器层的模型算法选择上, 考虑到在常见分类算法中, 随机森林(Random Forest, RF)^[27]和梯度提升决策树(Gradient Boosting Decision Tree, GBDT)^[28]有着出色的学习能力, 自适应提升(Adaptive Boosting, AdaBoost)^[29]考虑分类器权重使得其具有较高精度, K 最近邻(K-nearest Neighbors, KNN)^[30]精度高、对异常值不敏感、简单有效, 支持向量机(Support Vector Machine, SVM)^[31]泛化错误率低, 因此选择以上算法作为基学习器层的模型算法。除以上模型算法之外, 为满足 Stacking 对基学习器“多且不同”的要求, 也将常见的极限树(Extremely Tree, ET)和决策树(Decision Tree, DT)^[32]作为基学习器层的模型算法。

第二层应当选择泛化能力较强的模型, 能够实现从中归纳并纠正第一层多个基学习器对于训练模型的偏执情况^[33], 通过对第一层基模型产生的数据进行二次预测, 从而进一步提高预测表现。XGBoost^[34]是 GBDT 的一种高效实现, 其中的正则项能够防止过拟合, 并且采用并行优化方式, 考虑了训练数据为稀疏值的情况, 可以为缺失值或指定值指定分支的默认方向, 能大大提升算法的效率, 因此本文选择 XGBoost 算法作为元学习器。

3.2 Stacking 集成

本文构建的 Stacking 集成模型的原理如下:

1) 将原始样本数据划分为训练集与测试集。对训练集进行 N 折划分, 作为基学习器层中每个基模型输入的基础数据;

2) 针对各个基模型, 分别进行 N 次训练, 每次训练保持 $1/N$ 的样本作为基模型的验证数据, 其余的 $(1-1/N)$ 作为训练数据, 一个模型会对应 N 个预测结果, 再将 N 个预测结果进行加权平均;

3) 得到基模型运行 N 次的平均值之后, 拼接每一基模型的预测结果代入下一层;

4) 将各个基模型得到的预测结果, 拼接上各个样本真实的标签, 代入第二层模型进行训练, 最终得到的预测结果便是 Stacking 集成的最终预测结果。

本文的 Stacking 集成方法实现过程如图 3 所示。

4 评价指标

4.1 结果评价指标

混淆矩阵^[26]是衡量分类模型结果中最基本最直观的一种方式, 从混淆矩阵中可以对预测结果进行直观地了解, 如表 2 所示。

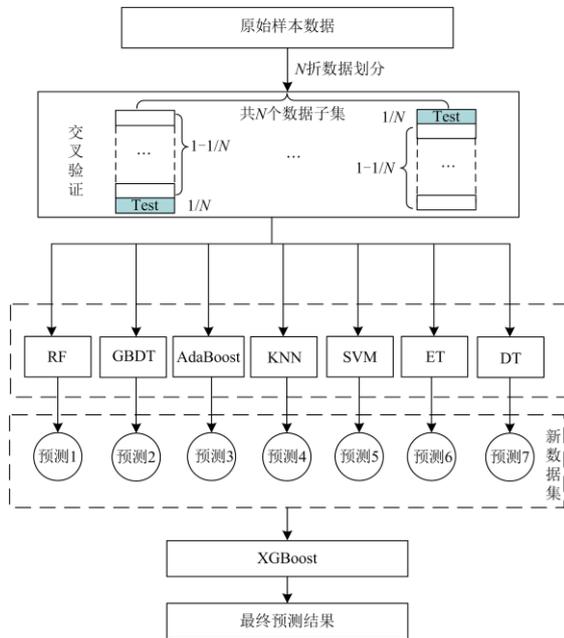


图 3 Stacking 集成过程示意图

Fig. 3 Diagram of Stacking integration process

表 2 混淆矩阵

Table 2 Confusion matrix

预测值	真实值	
	真实值 0	真实值 1
预测值 0	N_{TN}	N_{FN}
预测值 1	N_{FP}	N_{TP}

表 2 中, N_{TN} 表示预测和真实都为 0 的数目; N_{FP} 表示预测为 1 但是真实为 0 的数目; N_{FN} 表示预测为 0 但是真实为 1 的数目; N_{TP} 表示预测和真实值都为 1 的数目。

基于混淆矩阵, 准确率 A (Accuracy)、精确率 P (Precision) 与召回率 R (Recall) 是机器学习分类问题常用的评价指标^[26]。准确率 A 指所有被正确预测的样本数(包括正确预测正类和正确预测负类)占总样本数的比例。精确率 P 表示预测值为 1 的样本中真实值为 1 的样本占比。召回率 R 表示真实值为 1 的样本有多少被正确预测。准确率 A 、精确率 P 、召回率 R 的计算公式如式(3)一式(5)所示。

$$A = \frac{N_{TN} + N_{TP}}{N_{TN} + N_{TP} + N_{FN} + N_{FP}} \quad (3)$$

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (4)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (5)$$

由于召回率 R 与精确率 P 往往是此消彼长的,

单取某一个指标作为评判标准过于片面。因此需要提出一个新指标来综合两个指标进行评判。

F 分数(F -Score)^[26]是综合考虑两个指标的新指标, 其数学定义如式(6)所示。

$$F = \frac{(1 + \gamma^2)}{\gamma^2 \times P + R} \times P \times R \quad (6)$$

参数 γ 用于衡量 P 和 R 在所研究问题中的相对重要程度。当 $\gamma < 1$ 时表明精确率 P 更重要; $\gamma > 1$ 时表明召回率 R 更重要; $\gamma = 1$ 时表明召回率和精确率同样重要。由于本文研究的问题是台风灾害下的停电网格分布预测, 在电网公司制定灾害应急策略时, 在该问题上对已停电网格的关注度远远高于未停电网格, 因此本问题更加重视召回率的提高而对精确率可以采取一定的牺牲, 本文取 $\gamma = 3$ ^[17]。当两个模型的 F 分数相差不大时, 则需要根据实际研究问题关于 P 和 R 的具体要求来对最终的模型进行评价。

4.2 模型评价指标

ROC(Receiver Operating Characteristic)曲线与 AUC(Area Under Curve)值是判断二分类预测模型优劣比较直观的指标^[26]。AUC 值即 ROC 曲线下的面积。模型的 AUC 取值不同, 对模型的优劣评价也不同。

- 1) $AUC < 0.5$, 说明模型比随机猜测还差;
- 2) $AUC = 0.5$, 模型跟随机猜测一样, 模型没有预测价值;
- 3) $0.5 < AUC < 1$ 时, 模型优于随机猜测, 具有一定的预测价值, 且越接近 1 分类器性能越好;
- 4) $AUC = 1$ 时, 模型具有完美的预测功能。

5 算例仿真

5.1 仿真数据

本文以广东省某县为研究对象, 数据选取 2014 年第 9 号台风“威马逊”、第 15 号台风“海鸥”以及 2015 年第 22 号超强台风“彩虹”共三个台风的数据作为样本数据。三个台风在历史上都对该县造成停电影响。

数据预处理部分采用上文所述方法进行处理, 然后选取“威马逊”及“海鸥”两个台风的数据作为训练集, 台风“彩虹”的数据用于对模型预测效果的测试。

5.2 仿真结果与分析

模型在训练集上的预测结果如表 3 所示。

从表 3 可以看出, 模型训练结果指标几乎都达到了 0.8 或以上, 总体表现良好。利用混淆矩阵来

表示本文模型在该县总共 1 641 个网格台风“彩虹”下的整体预测结果,如表 4 所示。

表 3 训练集预测结果

Table 3 Prediction results of training set

准确率 A	精确率 P	召回率 R	F 分数
0.824 4	0.813 0	0.798 8	0.800 2

表 4 仿真结果混淆矩阵

Table 4 Simulation result confusion matrix

预测值	真实值	
	真实值 0	真实值 1
预测值 0	936	32
预测值 1	333	340

仿真得到在台风“彩虹”上的预测效果,各项评价指标如表 5 所示。

表 5 仿真结果各项指标

Table 5 Various indicators of simulation results

准确率 A	精确率 P	召回率 R	F 分数
0.777 6	0.505 2	0.914 0	0.845 6

结合表 4 和表 5 可以看出:在所有 1 641 个网格中,准确预测网格有 $936+340=1276$ 个,准确率达到 0.777 6;在所有真实发生停电的 $32+340=372$ 个网格中,预测为停电的网格数达 340 个,只有 32 个网格被错判,召回率达到 0.914 0, F 分数达到了 0.845 6,但是精确率偏低只有 0.505 2。说明模型能够很好地识别已停电网格。在总的网格层面,准确率较好,但是精确率较低,经多次调试后,发现模型的精确率仍然不高,其原因可能是考虑的特征不全面。如在 2014 年的台风过后,网格 A 中倒塌的杆塔更换为新杆塔,该网格的“抗风保电”水平相比原来有所提升,但是在特征上,仅考虑了杆塔数量以及拉线加固效应,不能够反映这一改变。在 2015 年的数据预测下,网格 A 可能因为忽视“抗风保电”水平的提升而使得模型将其预测为停电网格,从而导致精确率不高。因此未来的研究需进一步细化考虑停电影响因素,扩充更多特征。

为了对预测结果进行更直观地体现,为后期电网公司制定应急策略提供参考,将实际停电网格与预测结果在 ArcGIS 10.7 软件上进行可视化,实际停电网格分布与预测停电空间分布对比如图 4 和图 5 所示。

从图 4 与图 5 可以看出,实际停电空间分布大致可分左中右三部分,而预测结果与实际停电空间左中右的分布趋势大致相同,几乎包含了所有真实停电的网格,但同时停电网格周围也有未停电的网格被预测为停电网格。

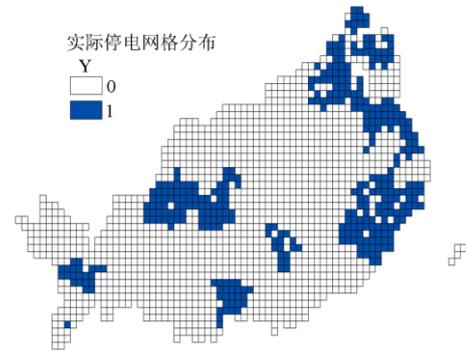


图 4 实际停电空间分布

Fig. 4 Actual power outage spatial distribution

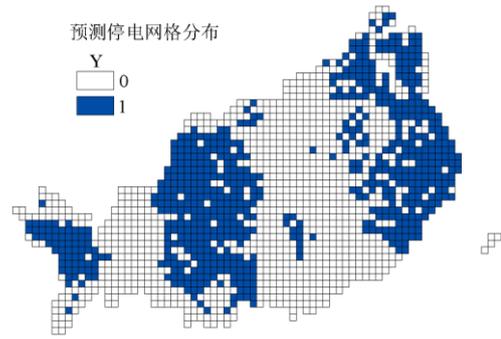


图 5 预测停电空间分布

Fig. 5 Predicted spatial distribution of outages

仿真得到模型的 ROC 曲线如图 6 所示。

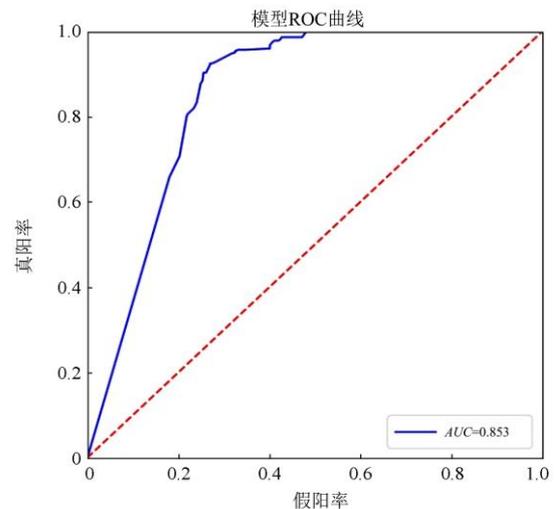


图 6 模型的 ROC 曲线

Fig. 6 ROC curve of model

从图 6 的 ROC 曲线可以看出,其 AUC 的值达到了 0.853,根据 ROC 曲线特性,当 $0.7 < AUC < 0.9$ 时有一定准确性,表明自变量与因变量的关联紧密,

同时说明网格停电情况与所选的评价指标具有良好的相关性,证明本文方法比随机预测要好,说明了本文方法是有效的,具有一定的预测价值。

将 Stacking 集成模型预测结果指标与各单一算法预测指标进行比较,如图 7 所示。

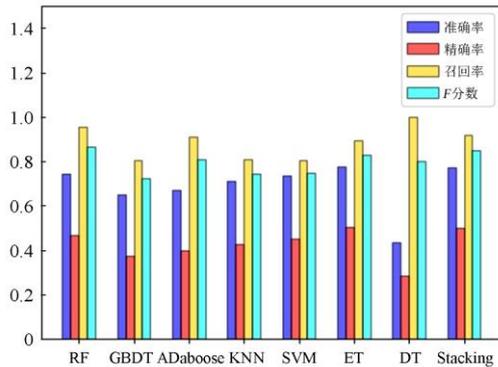


图 7 Stacking 集成与其他单一模型比较

Fig. 7 Stacking integration compared to other single models

从图 7 中可以看到,总体上 Stacking 集成的结果比单一算法得到的结果好,具体表现为:模型准确率、精确率与召回率都排在前列,图 7 中,虽然 RF 的召回率(0.956 0)与 DT 算法的召回率(1.000 0)指标略高于 Stacking 集成模型的召回率(0.913 9)指标,但是在 RF 与 DT 算法在精确率(0.467 8、0.285 9)与准确率(0.743 4、0.433 8)都比 Stacking 集成模型低, F 分数与 Stacking 集成模型相差不大。因此综合考虑精确率、准确率、召回率来说,Stacking 集成模型是最优的,验证了 Stacking 集成模型具有算法融合并在融合的基础上实现进一步改进的功能。

6 结论

1) 本文基于机器学习中的 Stacking 思想,提出了一种用于台风灾害下停电空间预测的方法。仿真结果表明本文方法在停电空间分布的趋势上与实际停电网格分布趋势较为一致,验证了本文模型在停电预测上的实际应用价值。

2) 本文方法在停电预测上的预测准确率达到 0.777 6,在对实际已停电网格预测方面,成功预测为停电网格的比例(召回率)达到了 0.914 0,说明了本文方法在预测停电方面的可行性及有效性。

3) 通过对比预测结果表明,Stacking 集成通过对多个模型的集成融合,对单一模型的预测性能具有一定的提升。

4) 利用已有数据,扩大停电预测研究区域,进一步验证模型在不同区域的适用性,并根据现有电

网系统灾害下的实时数据监测接收技术,而停电预测的实时更新功能将是下一步的重点研究方向。

参考文献

- [1] 常康,徐泰山,郁琛,等. 自然灾害下电网运行风险控制策略探讨[J]. 电力系统保护与控制, 2019, 47(10): 73-81.
CHANG Kang, XU Taishan, YU Chen, et al. Discussion of power system operation risk control strategy in natural disasters[J]. Power System Protection and Control, 2019, 47(10): 73-81.
- [2] 黄伟,黄廷城,王立勇,等. 基于态势感知的电网台风预警防御框架研究综述[J]. 电力系统保护与控制, 2018, 46(11): 162-169.
HUANG Wei, HUANG Tingcheng, WANG Liyong, et al. Research on typhoon early warning and defense framework of power grid based on situation awareness[J]. Power System Protection and Control, 2018, 46(11): 162-189.
- [3] D'AMICO D F, QUIRING S M, MADERIA C M, et al. Improving the hurricane outage prediction model by including tree species[J]. Climate Risk Management, 2019, 25: 1-15.
- [4] 黄勇,魏瑞增,周恩泽,等. 台风灾害下输电线路损毁预警方法[J]. 电力系统自动化, 2018, 42(23): 142-147.
HUANG Yong, WEI Ruizeng, ZHOU Enze, et al. Early warning of transmission line damage under typhoon disaster[J]. Automation of Electric Power Systems, 2018, 42(23): 142-147.
- [5] 尹超雄,唐武勤,温灵锋,等. 台风天气下配电网可靠性的新型评估算法[J]. 电力系统保护与控制, 2018, 46(4): 138-143.
YIN Chaoxiong, TANG Wuqin, WEN Lingfeng, et al. A new method for reliability evaluation of distribution network considering the influence of typhoon[J]. Power System Protection and Control, 2018, 46(4): 138-143.
- [6] 王永明,殷自力,李琳,等. 台风灾害场景下考虑运行状态的配电网风险评估方法[J]. 电力系统及其自动化学报, 2018, 30(12): 60-65.
WANG Yongming, YIN Zili, LI Lin, et al. Risk assessment method for distribution network considering operation state in the scene of typhoon disaster[J]. Proceedings of the CSU-EPSA, 2018, 30(12): 60-65.
- [7] 朱凌,陈涛威,周晨,等. 考虑风速风向联合分布的大风灾害下电力断线倒塔概率预测[J]. 电力系统保护与控制, 2019, 47(2): 115-122.
ZHU Ling, CHEN Taowei, ZHOU Chen, et al. Probability prediction of transmission line breakage and tower topple over under wind disaster considering the joint distribution

- of wind speed and wind direction[J]. *Power System Protection and Control*, 2019, 47(2): 115-122.
- [8] 侯慧, 俞菊芳, 黄勇, 等. 台风侵袭下输电线路风偏跳闸风险评估[J]. *高电压技术*, 2019, 45(12): 3907-3915.
HOU Hui, YU Jufang, HUANG Yong, et al. Risk assessment of windage yaw trip of transmission lines under the attack of typhoon[J]. *High Voltage Engineering*, 2019, 45(12): 3907-3915.
- [9] 陈莹, 王松岩, 陈彬, 等. 台风环境下考虑地理高程信息的输电通道结构失效故障概率评估方法[J]. *电网技术*, 2018, 42(7): 2295-2302.
CHEN Ying, WANG Songyan, CHEN Bin, et al. Evaluation of the failure probability of power transmission corridors during typhoons using digital elevation information[J]. *Power System Technology*, 2018, 42(7): 2295-2302.
- [10] 包博, 程韧俐, 熊小伏, 等. 一种计及微地形修正的输电线路台风风险预警方法[J]. *电力系统保护与控制*, 2014, 42(14): 79-86.
BAO Bo, CHENG Renli, XIONG Xiaofu, et al. A typhoon risk early warning method for power transmission line considering micro-terrain correction[J]. *Power System Protection and Control*, 2014, 42(14): 79-86.
- [11] 王泽斌, 王松岩, 陈莹, 等. 强台风环境下考虑微地形因素的输电通道结构安全概率评估方法[J]. *电力自动化设备*, 2020, 40(1): 184-191.
WANG Zebin, WANG Songyan, CHEN Ying, et al. Safety probability assessment method considering micro-topography for transmission grid under strong typhoon environment[J]. *Electric Power Automation Equipment*, 2020, 40(1): 184-191.
- [12] 李彦斌, 韩颖, 张嵘, 等. 气象因素对电力安全事故影响的模型[J]. *电网技术*, 2013, 37(6): 1683-1687.
LI Yanbin, HAN Ying, ZHANG Rong, et al. Research on impact model of meteorological factors on the power accidents[J]. *Power System Technology*, 2013, 37(6): 1683-1687.
- [13] 张勇军, 陈超, 许亮. 基于模糊聚类和相似度的电力系统可靠性原始参数预估[J]. *电力系统保护与控制*, 2011, 39(8): 1-5.
ZHANG Yongjun, CHEN Chao, XU Liang. Prediction of original reliability parameters of power system based on fuzzy clustering and similarity[J]. *Power System Protection and Control*, 2011, 39(8): 1-5.
- [14] LIU Haibin, DAVIDSON R A, ROSOWSKY D V, et al. Negative binomial regression of electric power outages in hurricanes[J]. *Journal of Infrastructure Systems*, 2005, 11(4): 258-267.
- [15] LIU Haibin, DAVIDSON R A, APANASOVICH T V. Spatial generalized linear mixed models of electric power outages due to hurricanes and ice storms[J]. *Reliability Engineering & System Safety*, 2008, 93(6): 897-912.
- [16] HAN S R, GUIKEMA S D, QUIRING S M. Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models[J]. *Risk Analysis*, 2009, 29(10): 1443-1453.
- [17] 侯慧, 耿浩, 肖祥, 等. 台风灾害下用户停电区域预测及评估[J]. *电网技术*, 2019, 43(6): 1948-1954.
HOU Hui, GENG Hao, XIAO Xiang, et al. Research on prediction and evaluation of user power outage area under typhoon disaster[J]. *Power System Technology*, 2019, 43(6): 1948-1954.
- [18] YUAN Shanshui, QUIRING S M, ZHU Ling, et al. Development of a typhoon power outage model in Guangdong, China[J]. *International Journal of Electrical Power & Energy Systems*, 2020, 117: 1-12.
- [19] WANIK D W, ANAGNOSTOU E N, HARTMAN B M, et al. Storm outage modeling for an electric distribution network in Northeastern USA[J]. *Natural Hazards*, 2015, 79(2): 1359-1384.
- [20] 叶圣永, 王晓茹, 刘志刚, 等. 基于 Stacking 元学习策略的电力系统暂态稳定评估[J]. *电力系统保护与控制*, 2011, 39(6): 12-16, 23.
YE Shengyong, WANG Xiaoru, LIU Zhigang, et al. Power system transient stability assessment based on Stacking meta-learning strategy[J]. *Power System Protection and Control*, 2011, 39(6): 12-16, 23.
- [21] 邓威, 郭钊秀, 李勇, 等. 基于特征选择和 Stacking 集成学习的配电网网损预测[J]. *电力系统保护与控制*, 2020, 48(15): 108-115.
DENG Wei, GUO Yixiu, LI Yong, et al. Power losses prediction based on feature selection and Stacking integrated learning[J]. *Power System Protection and Control*, 2020, 48(15): 108-115.
- [22] 程超鹏, 彭显刚, 曾勇斌, 等. 相异模型下 Stacking 集成结构的异常用电用户识别方法[J]. *电网技术*, 2021, 45(12): 4828-4836.
CHENG Chaopeng, PENG Xiangang, ZENG Yongbin, et al. An abnormal power user recognition method for Stacking integrated structures with different models[J]. *Power System Technology*, 2021, 45(12): 4828-4836.
- [23] 吴建华. ArcGIS 软件与应用[M]. 北京: 电子工业出版社, 2017.
- [24] 钟继. 地形与地表类型对台风近地风场多尺度模拟的影响研究[D]. 深圳: 哈尔滨工业大学(深圳), 2019.
ZHONG Ji. The impact of topography and land use on multi-scale simulation for near ground wind field of

typhoon[D]. Shenzhen: Harbin Institute of Technology (Shenzhen), 2019.

[25] 赵晨光, 周次明, 庞彦东, 等. 基于独热码有限状态机的斐索干涉解调相位补偿方法[J]. 光子学报, 2020, 49(5): 1-9.
ZHAO Chenguang, ZHOU Ciming, PANG Yandong, et al. Phase compensation method of Fizeau interference demodulation based on one-hot encoded finite state machine[J]. Acta Photonica Sinica, 2020, 49(5): 1-9.

[26] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.

[27] 李焱, 贾雅君, 李磊, 等. 基于随机森林算法的短期电力负荷预测[J]. 电力系统保护与控制, 2020, 48(21): 117-124.
LI Yan, JIA Yajun, LI Lei, et al. Short term power load forecasting based on a stochastic forest algorithm[J]. Power System Protection and Control, 2020, 48(21): 117-124.

[28] YANG Jinshan, ZHAO Chenyue, YU Haotong, et al. Use GBDT to predict the stock market[J]. Procedia Computer Science, 2020, 174: 161-171.

[29] 游文霞, 申坤, 杨楠, 等. 基于 AdaBoost 集成学习的窃电检测研究[J]. 电力系统保护与控制, 2020, 48(19): 151-159.
YOU Wenxia, SHEN Kun, YANG Nan, et al. Research on electricity theft detection based on AdaBoost ensemble learning[J]. Power System Protection and Control, 2020, 48(19): 151-159.

[30] 陈招安, 黄纯, 张志丹, 等. 基于 T 型灰色关联度和 KNN 算法的低压配电网台区拓扑识别方法[J]. 电力系统保护与控制, 2021, 49(1): 163-169.
CHEN Zhaoan, HUANG Chun, ZHANG Zhidan, et al. Topology identification method of a low voltage distribution network based on T-type grey correlation degree and KNN algorithm[J]. Power System Protection and Control, 2021, 49(1): 163-169.

[31] HOU Kaiyuan, SHAO Guanghui, WANG Haiming, et al. Research on practical power system stability analysis algorithm based on modified SVM[J]. Protection and Control of Modern Power Systems, 2018, 3(2): 119-125.

[32] AHMAD M W, REYNOLDS J, REZGUI Y. Predictive modelling for solar thermal energy systems: a comparison of support vector regression, random forest, extra trees and regression trees[J]. Journal of Cleaner Production, 2018, 203: 810-821.

[33] 史佳琪, 张建华. 基于多模型融合 Stacking 集成学习方式的负荷预测方法[J]. 中国电机工程学报, 2019, 39(14): 4032-4041.
SHI Jiaqi, ZHANG Jianhua. Load forecasting based on multi-model by Stacking ensemble learning[J]. Proceedings of the CSEE, 2019, 39(14): 4032-4041.

[34] 唐雅洁, 林达, 倪筹帷, 等. 基于 XGBoost 的双层协同实时校正超短期光伏预测[J]. 电力系统自动化, 2021, 45(7): 18-27.
TANG Yajie, LIN Da, NI Chouwei, et al. XGBoost based bi-layer collaborative real-time calibration for ultra-short-term photovoltaic prediction[J]. Automation of Electric Power Systems, 2021, 45(7): 18-27.

收稿日期: 2021-05-17; 修回日期: 2021-07-14

作者简介:

侯 慧(1981—), 女, 博士, 副教授, 博士生导师, 研究方向为电力系统风险评估、能源互联网、电动汽车与电网互动等; E-mail: houhui@whut.edu.cn

陈 希(1997—), 男, 通信作者, 硕士研究生, 研究方向为电力系统风险评估. E-mail: 1763854410@qq.com

(编辑 许 威)