文章编号:1001-2060(2021)05-0073-09

基于 Stacking 算法集成模型的电厂 NO_x 排放预测

李 阳[1],黄 伟[1],席建忠[2]
(1. 上海电力大学 自动化工程学院,上海 200090;2. 杭州华电江东热电有限公司,浙江 杭州 311228)

摘 要:提出一种基于 Stacking 算法集成模型的 NO_x 排放预测方法。考虑不同算法的训练机理和观测角度,将门控循环单元(gated recurrent unit,GRU)、XGBoost(eXtreme gradient boosting)和随机森林(random forest,RF)等多个学习能力强、差异度大的模型进行融合,得到一个具有两层结构的集成模型,通过弹性网(elastic network,EN)克服 DCS 采集的数据集内存在的共线性和群组效应,然后构造特征变量作为集成模型的输入。以某电厂的历史运行数据进行测试,结果表明 Stacking 集成模型的预测均方误差为 6.945 mg/m^3,相比单模型降低了 13.350% ~ 52.186%,根据其准确的预测结果可以更好的调整设备运行参数,保证排放的污染物浓度控制在合适的范围内。

关 键 词:电厂;NO_x 排放;人工智能;stacking 算法集成模型

中图分类号:TP181 **文献标识码**:A DOI:10.16146/j.cnki.rndlgc.2021.05.012

[**引用本文格式**]李 阳,黄 伟,席建忠. 基于 Stacking 算法集成模型的电厂 NO_x 排放预测[J]. 热能动力工程,2021,36(5):73-81. LI Yang,HUANG Wei,XI Jian-zhong. NO_x emission forecasting based on stacking ensemble model[J]. Journal of Engineering for Thermal Energy and Power,2021,36(5):73-81.

NO_x Emission Forecasting based on Stacking Ensemble Model

LI Yang[1], HUANG Wei[1], XI Jian-zhong[2]
(1. School of Automation Engineering, Shanghai University of Electric Power, Shanghai, China, Post Code: 200090;
2. Hangzhou Huadian Jiangdong Thermal Power Co. Ltd., Hangzhou, China, Post Code: 311228)

Abstract: An stacking ensemble model forecasting NO_x emission was proposed. Considering the difference of data observation and training principles, a two-layer stacking model embedded various machine learning algorithms such as gated recurrent unit (GRU), eXtreme gradient boosting (XGBoost) and random forest (RF) was established to realize the NO_x emission prediction. The elastic network (EN) was applied to extract the data from the DCS and eliminate the coupling of each feature variable, and the extracted data was used as input of stacking ensemble model. With the data of thermal power plant history as practical examples, the results show that the root mean square error of stacking ensemble model is 6.945 mg/m^3, which is 13.350% ~ 52.186% lower than that of single models. The operation parameters of the equipment can be better adjusted according to accurate prediction results of stacking ensemble model, which is of great significance to ensure the NOx concentration in the appropriate range.

Key words: power plant, NO_x emission, artificial intelligence, stacking ensemble model

收稿日期:2020-07-14; **修订日期**:2020-08-05

基金项目:国家自然科学基金(61503237);上海市科委发电过程智能管控工程技术研究中心基金资助项目(14DZ2251100)

Fund-supported Project: National Natural Science Foundation of China (61503237); Shanghai Science and Technology Commission Power Generation Process Intelligent Control Engineering Technology Research Center (14DZ2251100)

作者简介:李 阳(1996-),男,山西忻州人,上海电力大学硕士研究生.

通讯作者:黄 伟(1966-),女,上海人,上海电力大学副教授.

引 言

目前,降低 NO_x 排放质量浓度的主要方法是对燃烧过程进行合理的控制以及对烟气进行脱销处理[1-4],而该技术实施的前提是建立一个能够反映 NO_x 与可控变量之间关系的模型[5]。考虑到 NO_x 生成机理十分复杂,受众多变量的影响,且各变量间存在复杂的耦合关系,因此很难得到准确的物理模型。

近年来随着计算机技术和人工智能的发展,对非线性问题的处理能力逐步提高,电站锅炉 NO_x 排放量的预测取得显著成果。文献[6]通过电站实时运行数据,建立了基于主成分分析方法和 RF 的 NO_x 排放模型,取得较精确的预测结果。文献[7]组合支持向量机(support vector machine,SVM)和 BP 网络建立 NO_x 排放预测模型,在锅炉实际燃烧优化中取得很好的效果。文献[8]利用互信息对众多运行参数进行筛选,得到了影响 NO_x 的主要变量,在降低模型复杂性的同时使模型的泛化能力得到增强。文献[9]采用最大互信息系数的变量选择方法和 NARX 神经网络对锅炉的 NO_x 排放特性进行建模,模型的泛化能力较强,但是准确率较差。

集成学习是机器学习的一种推广,是对一系列基模型进行训练,通过某种方法将各模型的输出结果进行整合处理,获得性能更好的集成模型[10-11]。文献[12-13]利用组合模型的方式进一步提高模型的精度,然而该方法不能充分发挥每个模型的优势。以 Bagging 和 Boosting 为代表的集成方法可以对训练效果差的样本赋值较高的权重进行二次学习,提高组合预测的泛化能力。然而该方法只能集成同类决策树模型,难以融合其他模型的优势特性,不能体现出不同算法间数据观测的差异性。

为此,将 GRU、XGBoost 等模型与集成学习相结合,提出一种 Stacking 集成学习架构下基于多个差异化模型的 NO_x 排放预测方法。首先,通过弹性网进行特征选择,建立随机变量序列来构建搜寻策略,准确地计算待选变量与主导变量间的相关度,将重要度高的属性作为模型的输入特征,然后根据 Stacking 集成原理以及相关算法的机理,考虑多种模型的数据观测空间,建立了多模型融合的 NO_x 预测模型,最后以某电厂锅炉为对象验证了该方法的有效性。

1 理论介绍

1.1 弹性网变量选择

弹性网(EN)是在最小绝对值收缩及变量选择方法(least absolute shrinkage and selection operator,LASSO)基础上的一种改进[14-15]。对于样本容量为 n,待选变量的个数为 p 的数据集 $(x_{i,j},y_i)(i=1,2,\cdots,n,j=1,2,\ldots,p)$,假设响应变量 y_i 已经中心化,预测变量 x_{ij} 标准化,即:

$$\sum_{i=1}^{n} y_i = 0 \tag{1}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{p} x_{i,j} = 0 \tag{2}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{p} x_{i,j}^2 = 1 \tag{3}$$

用回归模型表示:

$$y_i = \sum_{j=1}^{p}\beta_j x_{i,j} + \varepsilon_i \tag{4}$$

式中:ε_i —满足标准分布的误差项。

对于 β_j,弹性网回归系数的表达式为:

$$\beta_{EN} = \operatorname{argmin}\left(\|Y - X\beta\|^2 + \lambda_1\sum_{j=1}^{p}|\beta_j| + \lambda_2\sum_{j=1}^{p}\beta_j^2\right) \tag{5}$$

令 $\alpha = \dfrac{\lambda_1}{\lambda_1+\lambda_2}$,$\lambda = \lambda_1 + \lambda_2$,得到:

$$\beta_{EN} = \operatorname{argmin}\left(\|Y - X\beta\|^2 + \lambda\left[\alpha\sum_{j=1}^{p}|\beta_j| + (1-\alpha)\sum_{j=1}^{p}\beta_j^2\right]\right) \tag{6}$$

其中,$\alpha\sum_{j=1}^{p}|\beta_j| + (1-\alpha)\sum_{j=1}^{p}\beta_j^2$ 为弹性网的惩罚项。同时,定义相对约束[16]:

$$s = \frac{\sum_j|\hat{\beta}_j|}{\sum_j|\hat{\beta}_j^0|} \tag{7}$$

式中:$\hat{\beta}_j$ —回归参数的模型估计结果;$\hat{\beta}_j^0$ —回归参数的最小二乘估计值。

基于 EN 的特征选择算法步骤为:

步骤 1:将预测变量 $x_{i,j}$ 标准化,响应变量 y_i 中心化;

步骤 2:将待选变量 $x_{i,j}$ 的系数 β_j 初始化为零;

步骤 3:设置 L_2 范数的系数 λ_2 ;

步骤 4:在 $0 < s < 1 + \lambda_2$ 的范围内使用已选的变量进行预测并得到回归残差值,根据赤池信息量准则(akaike information criterion, AIC)判断是否筛选结束,其计算式为:

$$\mathrm{AIC} = p\lg\frac{1}{p}\sum_{i=1}^{p} r_i^2 + 2(k+1)^2 \tag{8}$$

式中:r —根据已选变量预测 y 对应的回归残差;k—已选特征的个数;p—全部变量的个数。

步骤 5:增加已选变量个数,直到 AIC 不再减小,根据 s 得出各变量的回归系数 β_j ,将此时 $\beta_j = 0$ 的变量剔除,完成特征变量的筛选。

1.2　Stacking 集成学习方式

Stacking 集成学习是先分别对第一层的基学习器进行训练,然后把每个基学习器的训练结果输入到第二层的元学习器中,最后由元学习器输出预测结果的一种集成方法[17],其结构如图 1 所示。

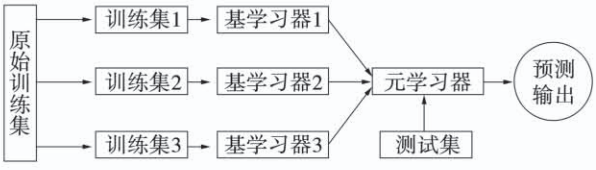

图 1　基于 Stacking 的集成学习方式

Fig. 1 Ensemble learning method based on Stacking

Stacking 集成学习的训练过程为[18]:对于训练集 $S = \{(x_i, y_i), i = 1,2,\cdots,N\}$ 中的第 i 个样本 (x_i, y_i),x_i 是样本的特征属性,y_i 是样本对应的预测值,K 为 x_i 的长度,即 $x_i = \{x_1, x_2, \cdots, x_K\}$ 。把 S 随机等分为 M 份,得到 $\{S_1\}, \{S_2\}, \cdots, \{S_M\}$,在第 m 次交叉训练过程中,定义 $\{S_m\}$ 为测试集,$\{S_{-m}\} = S - \{S_m\}$ 为训练集。第一层 M 个基学习器对训练集 $\{S_{-m}\}$ 学习后得到的基模型分别为 $L_1, L_2, \cdots, L_M$ 。M 个模型根据每一个样本 (x_i, y_i) 的特征 x_i 得到预测值为 $\bar{y}_{1,i}, \bar{y}_{2,i}, \cdots, \bar{y}_{M,i}$,与 y_i 组合得到 $z_i = \{y_i, \bar{y}_{1,i}, \bar{y}_{2,i}, \cdots, \bar{y}_{M,i}\}$,其中 $i = 1,2,\cdots,\frac{N}{M}$,把 $Z = \{z_1, z_2, \cdots, z_{\frac{N}{M}}\}$ 作为第二层元学习器的训练数据得到元模型 Y。Stacking 集成学习方式能够纠正第一层模型预测结果的误差,提升模型的性能[19]。

2　基于 Stacking 模型的 NO_x 预测方法

2.1　数据采集

锅炉燃烧受到多种因素的影响,每一台机组有上百个属性,其中,过半属性表示排气端特征、燃料情况、各阀门和断路器的开闭状态,而 NO_x 的生成主要取决于机组负荷、温度、压力、流量、煤量和风量等参数,因此需要将一些无影响或影响很小的变量剔除。

将 DCS 每隔 10 s 采集机组的运行参数分别用 $x_1, x_2, \cdots, x_{20}$ 来表示,用 y 来表示排出的 NO_x。由于不同属性值相差很大,并且具有不同的量纲,必须进行归一化操作以避免某些特征的数值太小而被淹没。采用标准化方法进行转换,即:

$$\tilde{x} = \frac{x - \bar{x}}{\sigma} \tag{9}$$

式中:$\bar{x}$ —原始属性 x 的平均值;σ —标准差;$\tilde{x}$ —标准化处理后的值。

2.2　特征选择

利用弹性网进行变量筛选的过程如图 2 和图 3 所示,图 2 为 AIC 随相对约束的变化趋势,图 3 为回归系数随相对约束的变化趋势。

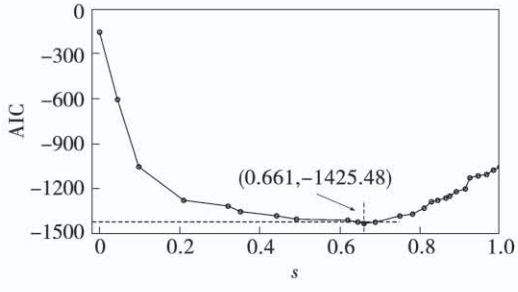

图 2　AIC 随相对约束的变化趋势

Fig. 2 Trend of AIC with relative constraints

当 L_2 范数 λ_2 =0.055 时,需要 26 步回归过程,当 $0 < s < 0.221$ 时,AIC 的值下降速度较快;在第 11 步时 s = 0.661, AIC 达到最小值 −1 425.48,此时 20 个待选变量的回归系数 β = [0, 0, 39.86, 0, 0, 0, 0, 0, 2.082, 4.71, 0, 12.96, 37.04, 33.53, 30.45, 24.93, 19.23, 0, 0, 0],即燃尽风 x_3 、二次风 x_9 、负

荷 x_{10}、一次风 x_{12}、烟气温度 x_{13}、烟气流量 x_{14}、炉膛压力 x_{15}、烟气含氧量 x_{16} 和总风量 x_{17} 共 9 个特征的系数不为零，但是当 $s>0.661$ 时 AIC 值缓慢上升，说明虽然后续选择的变量与 NO_x 之间有相关性，但在已经包含前 9 个变量的情况下，继续增加变量的个数会导致预测效果变差，因此将此时的 9 个属性作为预测 NO_x 的特征变量。

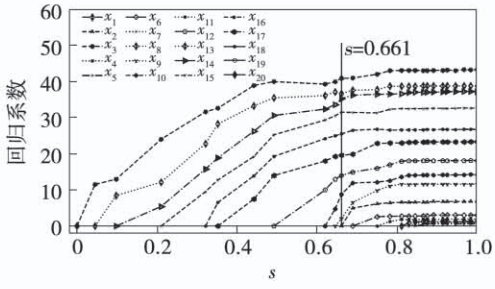

图 3　回归系数随相对约束的变化趋势

Fig. 3 Trend of regression coefficients with relative constraints

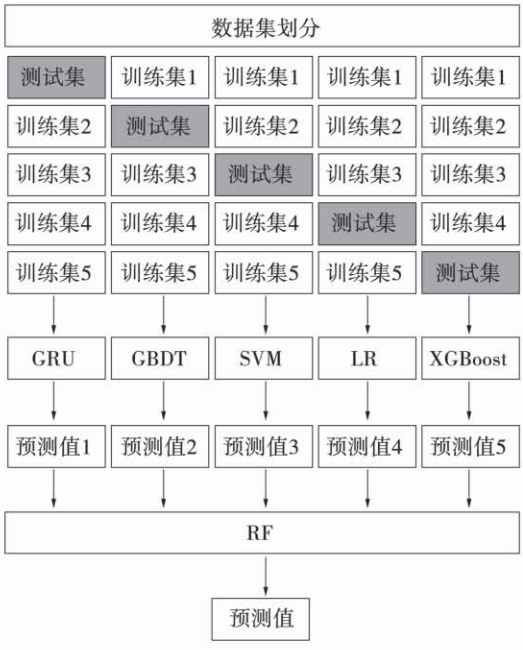

图 4　Stacking 集成模型结构

Fig. 4 Structure of stacking ensemble model

2.2　基学习器选择

在建立集成模型之前，既要分析每个单模型的预测能力，也要分析其组合效果。

在选择第一层的基模型时，鉴于不同单模型的预测能力，首先选择在时间序列预测领域具有良好表现的 GRU 深度学习模型和最强机器学习算法之一的 XGBoost，同时还要考虑其他性能优异的若干模型作为基学习器，因为学习能力强的基模型有助于整体预测效果的提升。RF 和梯度提升决策树(gradient boosting decision tree，GBDT)在各领域已经有着很广泛的应用，SVM 对于解决小样本、高维度的回归问题表现特有的优势，线性回归(linear regression，LR)因为理论成熟、训练高效也有着良好的实践应用效果。

第二层要选择泛化能力较强的模型，从中归纳出并纠正多个学习算法对于训练集的偏置情况，通过集合方式防止过拟合现象出现。因此，Stacking 集成模型第一层的基学习器初步选择 GRU、XGBoost、GBDT、LR 和 SVM，第二层选择 RF 作为元学习器，得到的模型架构如图 4 所示。

为了获得最优预测效果，在 Stacking 模型第一层需要选择差异度较大的模型作为基学习器。这是由于不同的算法本质是从不同的空间角度和数据角度来观测，再依据算法的观测状况及自身算法原理建立相应模型。因此，选择差异度较大的算法能够最大程度体现其优势，便于模型之间取长补短。为了体现不同模型的差异度，采用 Pearson 相关系数进行度量，以此分析基学习器之间的关联程度，Pearson 相关系数计算式为：

$$r_{x,y}=\frac{\sum_{i=1}^{m}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{m}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{m}(y_i-\bar{y})^2}} \tag{10}$$

为了避免数据被两层学习器重复学习而造成的过拟合现象，必须要对使用的数据进行合理的划分。在图 4 中根据所选择的 5 个基模型将原始数据集进行 5 次划分，对于每个基学习器，选择阴影区域的数据集作为测试集，其余的作为训练集，保证每个基模型互不重叠。每个基学习器对相应的测试集都有一组预测结果，5 组单模型的预测结果可合并为与原始训练集规模相同的新数据，然后再将其作为元学

习器的输入,训练得到 Stacking 模型,求得最终的预测值。

2.3　集成模型建立

建立的 NO_x 预测模型属于动态模型,即每个变量随时间变化而变化,特征变量均为按照时间索引的一系列数值,每个时刻的输出不仅受到当前时刻特征属性的影响,还与前 k 个时刻的输出有联系。因此,将原本属于时间序列的运行数据转换为适用于机器学习的"监督形式"的数据,转换后无量纲的数据如表 1 所示,实现了由输入特征到输出结果的变换,那么预测结果可以通过该函数表示:

$$y(t) = F(x_3(t-1),\cdots,x_{17}(t-1),y(t-1),\cdots,y(t-k)) \tag{11}$$

式中:F—Stacking 集成模型从原始数据中挖掘得到的隐含函数关系;$y(t)$—当前时刻的函数值,$x_3(t-1),\cdots x_{17}(t-1),y(t-1),\cdots,y(t-k)$—自变量,表示各属性在前一时刻的值。

表 1　监督学习数据

Tab. 1 Supervised learning data

$x_3(t-1)$	…	$x_{17}(t-1)$	$y(t-k)$	…	$y(t-1)$	$y(t)$
0.851	…	0.631	0.553	…	0.603	0.622
0.895	…	0.658	0.564	…	0.622	0.625
0.915	…	0.552	0.589	…	0.625	0.683
0.151	…	0.631	0.624	…	0.683	0.392
0.195	…	0.478	0.657	…	0.392	0.315
⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.791	…	0.539	0.754	…	0.728	0.705
0.748	…	0.527	0.740	…	0.705	0.682

基于 Stacking 集成模型的 NO_x 模型流程如图 5 所示。DCS 采集到的运行参数划分五次,每次均分为训练数据和测试数据,保证每次划分过程中测试数据完全不相同,对应图 4 中的阴影区域,从而对五个基学习器单独训练。根据 Stacking 集成理论融合多个差异性大、能力强的模型得到一个性能优秀的集成模型,将靠近待测点附近的数据作为集成模型的输入,输出未来一段时间内的 NO_x 变化趋势。

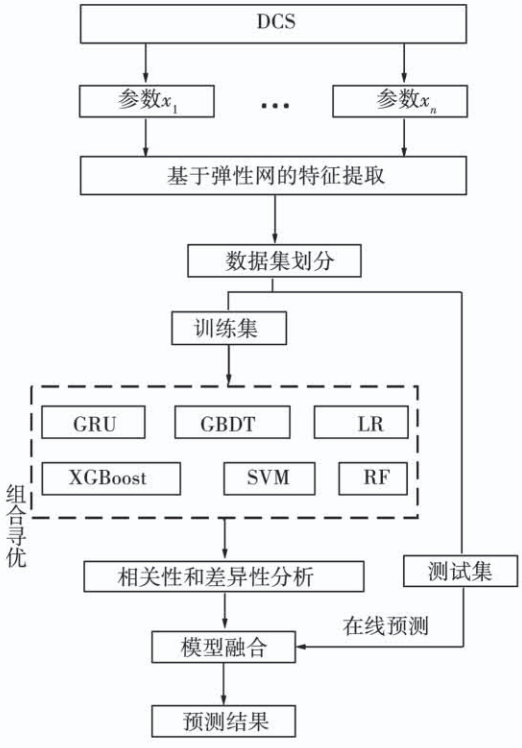

图 5　基于 Stacking 集成模型的 NO_x 排放预测流程

Fig. 5 Prediction process of NO_x with stacking model

3　算例分析

预测模型的评价指标选择平均相对误差 e_{MAPE}(mean absolute percentage error,MAPE)和相对均方误差 e_{RMSE}(root mean square error,RMSE):

$$e_{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{y_i} \tag{12}$$

$$e_{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{13}$$

式中:y_i 和 $\hat{y}_i$ —第 i 个测点的实际值和预测值;n —预测样本的数量。

3.1　单模型关联度计算

对于各基学习器,采用交叉验证的方法在训练集上使用不同的超参数训练,从而选择模型的最优超参数集。6 种单模型的超参数以及预测性能如表 2 所示。

由于 NO_x 当前时刻的值与前一个时刻之间存在一定的联系,也与过去的输入特征有关。GRU 可以有效利用累积信息,得到较好的预测结果;而 XG-

Boost 充分利用剩余的数据计算信息增益，保证了估计结果的准确性；GBDT 在梯度提升的过程中可以不断地学习前一次的结果，因此其效果要优于 RF 以及其他机器学习模型，体现了 Boosting 算法集成多个决策树的高泛化性。根据式(10)计算各模型预测结果的相关度，得到的相关系数值如图 6 所示。相关度越接近 1.0 颜色越深，反之相关度越靠近 0 对应的颜色越浅。

表 2 模型的超参数及预测结果

Tab. 2 Hyper-parameters and forecasting results of single models

单模型	超参数值	e_{MAPE}/%	e_{RMSE}/ mg · m^{-3}
GRU	神经元个数为 77 隐藏层数量为 3 优化器 = Adam	4.294	9.297
XGBoost	特征比例 = 0.452 学习率 = 0.0185 叶子数量 = 20 叶子最少单元 = 10	3.362	8.015
GBDT	决策树数量 = 160 最大特征比 = 0.208 最大深度 = 9	4.726	9.455
RF	决策树数量 = 118 叶子点最少样本 = 2 内部节点最小值 = 4	4.941	9.564
SVM	核参数 = 线性 标准差 = 0.318 惩罚系数 = 0.658	5.381	14.021
LR	布尔类型 = 真	5.856	14.525

由图 6 可知，各算法对应误差的关联度相对较高，其中，XGBoost、RF 和 GBDT 之间相关度最高，因为虽然三者原理有些许不同，但是均为基于决策树的集成算法，其数据观测方式的差异度很小。GRU、SVM 和 LR 在训练机理方面有着很大的差距，因此误差相关性也较低。

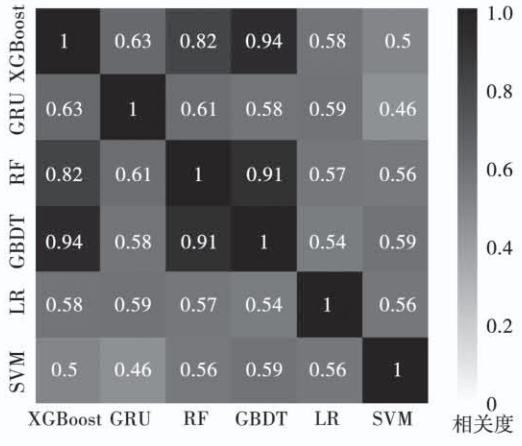

图 6 各模型预测结果的相关性

Fig. 6 Correlation of forecasting results for each model

3.2 Stacking 模型性能分析

考虑到 NO_x 排放量受季节影响，分别以四月和十月的数据进行试验，得到两组数据集 T1 和 T2，将设计的 Stacking 集成模型与 GRU、XGBoost 以及 LR 单模型进行对比，由于原始样本数据的采样间隔是 1 s，考虑到数据量庞大，需要对样本数据进行二次采样以避免数据太密集造成的信息冗余，采样间隔为 10 s。4 种方法在两组测试集上的预测结果如图 7 所示，横坐标表示按照二次采样间隔排序的预测时刻，纵坐标为此时 NO_x 的质量浓度。

从图 7 中可以看出，GRU 和 XGBoost 均能够很好地跟随 NO_x 质量浓度的变化趋势，但其预测的误差仍然较大，而 Stacking 模型在 NO_x 大幅度变化和平缓变化的时段可以更稳定地捕捉到其变化规律，这源于其充分发挥并融合了不同单模型的优势，拥有更好的鲁棒性和精准性。Stacking 模型的 e_{RMSE} 相比于 XGBoost，GRU，RF，GBDT，SVM 和 LR 分别降低了 25.298%，13.350%，26.547%，27.384%，50.467% 和 52.186%，e_{MAPE} 分别降低了 1.420%，0.488%，1.852%，2.067%，2.507% 和 2.982%。Stacking 模型在 e_{RMSE} 和 e_{MAPE} 指标上都有明显的下降，表明其预测效果最好。

3.3 不同基学习器组合方式对比

为验证上述单模型组合方式的合理性，表 3 给

出了不同组合方式下集成模型的预测误差。

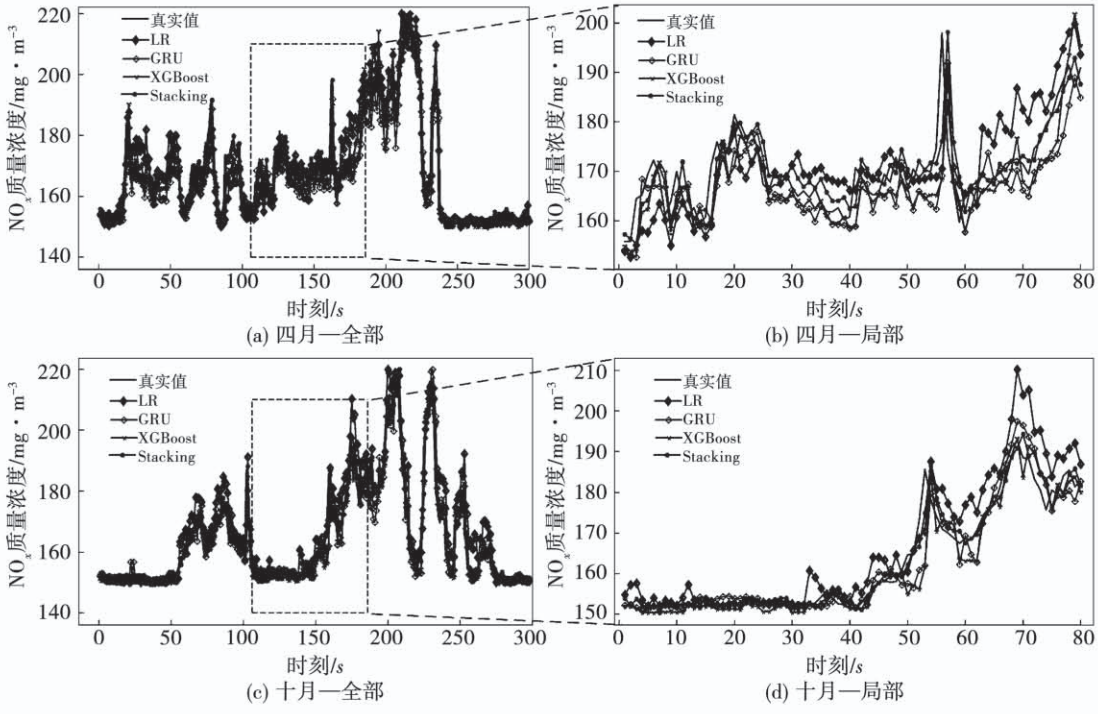

图 7　Stacking 与 LR、GRU、XGBoost 的预测结果对比

Fig. 7 Prediction with stacking, LR, GRU and XGBoost

根据表 3 可知,组合不同单模型得到的集成模型预测效果有明显的差距。按照方式 5 得到的集成模型的效果比单模型 LR 还要差。这是因为其第一层模型包含了相关度最大的 RF,GBDT 和 XGBoost,导致数据被反复训练,从而降低了模型的泛化性;组合方式为 1 和 3 的集成模型第一层基学习器也包含了基于决策树的模型,虽然均为学习能力较强的模型,但是模型间的相关度并不是最小,因此得到的集成模型和单模型的效果差别不大;组合方式为 2 和 4 的两种集成模型的性能要优于上述集成模型以及 6 种单模型,原因在于第一层单模型之间的相关性均为最小值,能够体现不同算法的优势,提高了融合模型的泛化性能;根据所提方法构建的 Stacking 模型以 2.874% 的 e_{MAPE} 和 6.945 mg/m^3 的 e_{RMSE} 验证了该方法的正确性和必要性,不仅要使第一层单模型之间的差异性最大,还要保证每个模型均为强学习器。

表 3　不同基模型组合后的预测误差

Tab. 3 Forecasting error of ensemble models based on different base-learner aggregation style

方式	基学习器	元学习器	e_{MAPE}/%	e_{RMSE}/mg·m^{-3}
1	GRU,RF,GBDT	XGBoost	5.042	9.424
2	LR,SVM,XGBoost	GBDT	3.226	7.854
3	LR,GRU,GBDT,RF	RF	5.165	9.959
4	GBDT,GRU,SVM	LR	3.135	7.590
5	XGBoost,LR,RF,GBDT	XGBoost	6.029	15.095
6	XGBoost,LR,GRU,SVM	RF	2.874	6.945

4　结　论

以某电厂锅炉为例,采用弹性网筛选出影响 NO_x 生成的主要影响变量,提出基于 Stacking 学习方式的集成模型 NO_x 排放预测方法,并对比了单模型的预测效果,得出结论:

(1) 在选择基模型时,要保证各单模型之间的

相关度尽可能小,避免“过度学习”的出现,还要保证每个模型均为强学习器,最终以 GRU,XGBoost,SVM 和 LR 作为第一层的基模型,RF 作为第二层的元模型,获得了性能最优的集成模型,e_{RMSE}为 6.945 mg/m^3。

(2) Stacking 集成的方式比较复杂,要对每个基模型进行训练,即使减少基学习器的个数,计算过程仍然非常耗时,后续的研究中要设置分布式计算环境,在保证精确度的同时提升模型的集成速度。

参考文献:

[1] 余廷芳,刘 冉. 基于 RBF 神经网络和 BP 神经网络的燃煤锅炉 NO_x 排放预测[J]. 热力发电,2016,45(8):94 - 98,113.
YU Ting-fang,LIU Ran. NO_x emission prediction of coal-fired boiler based on RBF neural network and BP neural network[J]. Thermal Power Generation,2016,45(8):94 - 98,113.

[2] 王文广,赵文杰. 基于 GRU 神经网络的燃煤电站 NO_x 排放预测模型[J]. 华北电力大学学报(自然科学版),2020,47(1):96 - 103.
WANG Wen-guang,ZHAO Wen-jie. A prediction model of NO_x emission from coal-fired power plants based on GRU neural network[J]. Journal of North China Electric Power University (Natural Science Edition),2020,47(1):96 - 103.

[3] 曹 勇,周 托,那永洁. 循环流化床锅炉炉膛低氧燃烧加尾部补燃降低 NO_x 排放的试验研究[J]. 锅炉技术,2019,50(2):35 - 42.
CAO Yong,ZHOU Tuo,NA Yong-jie. Experimental study on reducing NO_x emission by low oxygen combustion and tail supplementary combustion in circulating fluidized bed boiler[J]. Boiler Technology,2019,50(2):35 - 42

[4] 高常乐,司风琪,任少君. 基于 LSTM 的烟气 NO_x 浓度动态软测量模型[J]. 热能动力工程,2020,35(3):98 - 104.
GAO Chang-le,SI Feng-qi,REN Shao-jun. Dynamic soft sensing model of NO_x concentration in flue gas based on LSTM[J]. Journal of Engineering for Thermal Energy and Power,2020,35(3):98 - 104.

[5] 闫来清,董 泽,孟 磊. 基于 OSC-CKPLS 方法的 SCR 出口 NO_x 排放预测[J]. 锅炉技术,2020,51(2):7 - 13.
YAN Lai-qing,DONG Ze,MENG Lei. NO_x emission prediction of SCR outlet based on OSC-CKPLSmethod[J]. Boiler Technology,2020,51(2):7 - 13.

[6] 许 壮,康英伟. 基于随机森林的火电机组 SCR 脱硝反应器建模[J]. 动力工程学报,2020,40(6):486 - 491,501.
XU Zhuang,KANG Ying-wei. Modeling of SCR denitration reactor of thermal power unit based on random forest[J]. Journal of Chinese Society Power Engineering,2020,40(6):486 - 491,501.

[7] 李鹏辉,刘 冉,余廷芳. 基于支持向量机和 BP 神经网络的燃煤锅炉 NO_x 排放预测[J]. 热能动力工程,2016,31(10):104 - 108.
LI Peng-hui,LIU Ran,YU Ting-fang. NO_x emission prediction of coal-fired boiler based on support vector machine and BP neural network[J]. Journal of Engineering for Thermal Energy and Power,2016,31(10):104 - 108.

[8] 马 平,李 珍,梁 薇. 基于互信息的辅助变量筛选及在火电厂 NO_x 软测量模型中的应用[J]. 科学技术与工程,2017,17(22): 249 - 254.
MA Ping,LI Zhen,LIANG Wei. Auxiliary variable selection based on mutual information and its application in NO_x soft sensor model of thermal power plant[J]. Science Technology and Engineering,2017,17 (22): 249 - 254.

[9] 王文广,王 朔,赵文杰. 基于最大信息系数变量选择的电站锅炉 NO_x 排放量在线预估[J]. 华北电力大学学报(自然科学版),2019,46(6): 66 - 72.
WANG Wen-guang,WANG Shuo,ZHAO Wen-jie. On line prediction of NO_x emission of utility boiler based on the selection of maximum information coefficient variable[J]. Journal of North China Electric Power University (Natural Science Edition),2019,46(6):66 - 72.

[10] LI Zhi-yi,LIU Xuan,CHEN Li-yuan. Load interval forecasting methods based on an ensemble of Extreme Learning Machines[C]//Proceedings of 2015 IEEE Power & Energy Society General Meeting,Denver,2015.

[11] 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 - 4042.
SHI Jia-qi,ZHANG Jian-hua. Load forecasting method based on multi model fusion stacking integrated learning method[J]. Proceedings of the CSEE,2019,39(14):4032 - 4042.

[12] 崔建国,李慧华,于明月,等. 基于 LSSVM 与 WNN 的燃气轮机状态趋势预测[J]. 火力与指挥控制,2018,43(8):160 - 163,167.
CUI Jian-guo,LI Hui-hua,YU Ming-yue,et al. State trend prediction of gas turbine based on LSSVM and WNN[J]. Fire Control & Command Control,2018,43(8):160 - 163,167.

[13] 刘文霞,龙日尚,徐晓波,等. 考虑数据新鲜度和交叉熵的电动汽车短期充电负荷预测模型[J]. 电力系统自动化,2016,40(12):45 - 52.
LIU Wen-xia,LONG Ri-shang,XU Xiao-bo,et al. Forecasting model of short-term EV charging load based on data freshness and

cross entropy[J]. Automation of Electric Power Systems,2016,40(12):45 - 52.

[14] 葛　阳,郭兰中,牛曙光,等. 基于 t - SNE 和 LSTM 的旋转机械剩余寿命预测[J]. 振动与冲击,2020,39(7):223 - 231,273.
GE Yang,GUO Lan-zhong, NIU Shu-guang,et al. Prediction of residual life of rotating machinery based on t-sne and LSTM[J]. Vibration and shock,2020,39(7):223 - 231,273.

[15] 张淑清,杨振宁,张立国,等. 基于弹性网降维及花授粉算法优化 BP 神经网络的短期电力负荷预测[J]. 仪器仪表学报,2019,40(7):47 - 54.
ZHANG Shu-qing,YANG Zhen-ning,ZHANG Li-guo,et al. Short term power load forecasting based on elastic network dimensionality reduction and flower pollination algorithm optimization BP neural network[J]. Chinese Journal of Scientific Instrument,2019,40(7):47 - 54.

[16] 陈　磊,范　宏. 集成学习框架下的个人信用评分模型研究[J]. 中国市场,2020(20):164 - 168.
CHEN Lei, FAN Hong. Research on Personal Credit Scoring Model under the Framework of Integrated Learning[J]. Chinese Market,2020(20):164 - 168.

[17] 高铭远. 基于集成模型的生物医学命名实体识别研究[D]. 大连海事大学, 2020.
GAO Ming-yuan. Research on Biomedical Named Entity Recognition Based on Integrated Model[D]. Dalian Maritime University,2020.

[18] 张　展,韩　华,崔晓钰. 基于多种模型集成学习的制冷系统故障诊断[J]. 热能动力工程,2020,35(5):153 - 162.
ZHANG Zhan,HAN Hua,CUI Xiao-yu. Refrigeration system fault diagnosis based on integrated learning of multiple models[J]. Thermal Energy and Power Engineering, 2020, 35(5):153 - 162.

[19] 徐　萌,席泽西,王雍赟. 基于集成学习的航空发动机故障诊断方法[J]. 中国民航大学学报,2019,37(2):29 - 33,42.
XU Meng,XI Ze-xi,WANG Yong-yun. Aero engine fault diagnosis method based on ensemble learning[J]. Journal of Civil Aviation University of China,2019,37(2):29 - 33,42.

(金圣迪　编辑)