# **Short-Term Electricity Load Forecasting Based on the XGBoost Algorithm**

Guangye Li1, Wei Li1, Xiaolei Tian1, Yifeng Che2

Email: liguangye liaodian@163.com

Received: Aug. 5<sup>th</sup>, 2017; accepted: Aug. 19<sup>th</sup>, 2017; published: Aug. 24<sup>th</sup>, 2017

#### **Abstract**

High accurate forecasting of short-term power load is important to make reasonable production plan and scheduling task and achieve energy saving and economic operation of electric power management system (EMS). In this paper, the periodic similarity of power load fluctuation characteristics with different time scales is given. According to the historical data of the same time load of the previous day, the same time load of previous week and average load of the last 24 hours, meteorological data and the information about the working day and holiday event, a short-term power load extreme gradient boosting (XGBoost) model with multi-information fusion is built by using the gradient boosting algorithm. It can predict the trend of 24 o'clock power load ahead one day. The results, through a regional State Grid 24 o'clock curve forecast, show that the built XGBoost forecasting model has advantages over random forest, Bayesian and KNN methods in terms of speed and prediction accuracy.

#### Keywords

Short-Term Power Load Forecasting, Forecasting Accuracy, XGBoost, Predictor

# 基于XGBoost算法的短期电力负荷预报

李广野1,李 伟1,田小蕾1,车轶锋2

<sup>1</sup>国网辽宁省电力有限公司,辽宁 沈阳 <sup>2</sup>南京南瑞集团公司,江苏 南京

Email: liguangye\_liaodian@163.com

收稿日期: 2017年8月5日; 录用日期: 2017年8月19日; 发布日期: 2017年8月24日

## 摘要

短期电力负荷的精准预报是电力能源管理系统(EMS)合理安排生产调度计划、实现节能、经济运行的前

文章引用: 李广野, 李伟, 田小蕾, 车轶锋. 基于 XGBoost 算法的短期电力负荷预报[J]. 智能电网, 2017, 7(4): 274-285. DOI: 10.12677/sg.2017.74031

<sup>&</sup>lt;sup>1</sup>State Grid Liaoning Electric Power Co., Ltd., Shenyang Liaoning

<sup>&</sup>lt;sup>2</sup>Nanjing Nanrui Group Company, Nanjing Jiangsu

提条件和重要保障。本文针对电力负荷波动特征具有不同时间尺度的周期相似性,根据前一天同一时刻、前一周同一时刻负荷和最近24小时平均负荷历史数据、融合气象数据以及工作日、节假日时间事件信息,采用梯度提升算法建立多信息融合的短期电力负荷极限梯度提升(XGBoost)模型,提前一天预测24点电力负荷变化趋势。通过某地区电网日负荷24点曲线预报的实验结果,表明所构建的电力负荷XGBoost预报模型相比随机森林、贝叶斯和KNN方法在计算速度和预测精度方面具有优势。

## 关键词

短期负荷预测,预测精度,XGBoost,预测因子

Copyright © 2017 by authors and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY). http://creativecommons.org/licenses/by/4.0/



Open Access

#### 1. 引言

负荷预测是根据电力系统的运行特性、自然条件等诸多影响因素,基于电力负荷的过去和现在运行状况,推测未来某特定时刻数值[1]。精准的电力负荷预报是电力能源管理系统(EMS)合理安排生产调度计划、实现节能、经济运行的前提条件和重要保障。电力负荷预报要求预测值与实际负荷的偏差必须小于电网 AGC (Automatic Generation Control)调节容量。由于目前电网 AGC 调节容量一般较小,对电力负荷预测模型精度提出了更高的要求,因此建立高性能实时电力负荷预报模型一直是学术界和工业界研究的热点[2]。

国内外许多学者对此进行了大量、深入的研究,提出了许多有效算法。传统的短期负荷预测方法[2]有回归分析法[3] [4]、时间序列法、卡尔曼滤波、外推法、指数平滑法、灰色预测法。随着机器学习理论的发展,人工神经网络[5]、支持向量机[6]、相关向量机等方法越来越广泛的用于负荷预测并取得令人满意的效果[7] [8] [9]。文献[10]将局部加权线性回归预测算法和云计算 Mapreduce 模型相结合建立了基于云计算的局部加权线性回归模型进行海量数据的短期负荷预测。文献[11]提出了一种主成分分析(PCA)和 BP神经网络相结合的电力负荷预测方法。文献[12]针对 AR 无法接纳多源数据与 BP神经网络算法结果受相似日影响大等弊端,提出一种优势互补的 AR-ANN 算法。由于电网规模不断扩大,负荷类型日趋多样化,数据数量和类型也在快速增加,上述方法在处理大量数据时有局限性,复杂的计算又使其运行时间显得较长。

XGBoost (Extreme Gradient Boosting) [13]是一种通过一组弱分类器的迭代计算以实现准确分类效果的集成学习框架。XGBoost 能够自动利用 CPU 的多线程并行计算,在算法的实现上又增加了正则项,极大地提高了模型运算效率,而且也使模型的泛化能力得到提升,在分布式应用上也有明显优势。相比于传统 GBDT (Gradient Boost Decision Tree),XGBoost 算法上使用了 Taylor 二阶信息,实现上采用了并行/多核计算方式,训练收敛更快速、准确率更高[14]。目前,在店铺销售预测[15]、事件分类[16]和客户行为预测[17]等方面,XGBoost 已经得到广泛应用。

本文采用计算速度快、效果好的 XGBoost 算法构建电力负荷预测模型,预测因子包括干球温度、露点温度、时刻(小时)、星期几、是否为工作日、前一周同一时刻负荷、前一天同一时刻负荷、前 24 小时平均负荷,预报输出是未来一天 24 点电网负荷变化趋势。以某地区电力负荷预报为研究对象进行了实验验证,实验结果验证了本文所提方法的有效性。

# 2. 极限梯度提升模型原理

XGBoost 是一个设计高效、灵活并且可移植的最优分布式决策梯度提升库,通过一组弱分类器的迭代计算实现准确的分类效果。它实现了梯度提升框架下的机器学习算法。基分类器是分类和回归树 (CART)。通常一个基模型很难保障预测精度,常见的解决思路是采用树的集成模型,把多个树的预测综合在一起。假设有K个树,树集成模型为

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), \ f_k \in \mathscr{F}$$
 (1)

其中  $f_k$  是函数空间  $\mathscr T$  里面的一个函数,而  $\mathscr T$  是包含所有回归树的函数空间。树集成模型参数包括每棵树的结构以及叶的得分,可以简单地使用函数  $f_k$  作为参数,  $\Theta = \{f_1, f_2, \cdots, f_K\}$  。目标函数包含损失函数和正则化项

$$Obj = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k), \ f_k \in \mathcal{F}$$

$$(2)$$

其中,损失函数为平方误差  $l(y_i,\hat{y}_i) = (y_i - \hat{y}_i)^2$ ,正则化项 $\Omega$ 包括 L1和 L2 正则。优化目标函数实际上是求解 CART 树的结构和叶分数,立即训练所有的树是不容易的,这要比传统的优化问题困难得多。文献[18] 提出采用梯度方式,加法策略,即固定已经学到的树,保留原来的模型不变,一次添加一个新的树到模型中。一般假设模型预报初始化  $\hat{y}_i^{(0)} = 0$ ,每次添加一个新的函数(树),迭代计算第 t 轮模型输出预报值  $\hat{y}_i^{(t)}$ 。

$$\hat{y}_{i}^{(t)} = \sum_{k=1}^{t} f_{k}(x_{i}) = \hat{y}_{i}^{(t-1)} + f_{t}(x_{i})$$
(3)

其中, $\hat{y}_i^{(t)}$ 表示第t轮的模型预测, $\hat{y}_i^{(t-1)}$ 表示保留前面t-1轮的模型预测, $f_i(x_i)$ 表示加入一个新的函数。在实验中通常用式(4)

$$\hat{\mathbf{y}}^{(t)} = \hat{\mathbf{y}}^{(t-1)} + \varepsilon f_t(\mathbf{x}_i) \tag{4}$$

来计算  $\hat{y}_{i}^{(r)}$ 。其中, $\varepsilon$  表示步长或者收缩,这意味着我们不会在每个步骤中进行完全优化预留未来几轮的机会,有助于防止过度拟合。模型预报式(3)代入目标函数(2)式中,得到

$$Obj^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) + C$$
(5)

根据 Taylor 二阶展式, 简化目标函数中损失函数, 重新定义目标函数:

$$Obj^{(t)} \approx \sum_{i=1}^{n} \left[ l\left(y_{i}, y_{i}^{(t-1)}\right) + g_{i} f_{t}\left(x_{i}\right) + \frac{1}{2} h_{i} f_{t}^{2}\left(x_{i}\right) \right] + \Omega\left(f_{t}\right) + C$$
(6)

其中,  $g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ ,  $h_i = \partial_{\hat{y}^{(t-1)}}^2 l(y_i, \hat{y}^{(t-1)})$ , C 为常数。移除常数项后,最终的目标函数只依赖于每个样本点在误差函数上的一阶  $g_i$  和二阶导数  $h_i$ 。

$$Obj^{(t)} \approx \sum_{i=1}^{n} \left[ g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t)$$
(7)

假设  $f_t$  为决策树,叶结点数目为T ,每个叶结点权重为 $\omega$  。决策树的学习过程,就是从根结点到叶结点细化的过程,树可以拆分成树结构部分 q 和叶权重部分 $\omega$  。结构函数 q 把输入样本 x 映射到叶结点 q 中,而  $\omega$  给定了叶权重。

$$f_t(x) = \omega_{q(x)}, \omega \in \mathbf{R}^T, q : \mathbf{R}^d \to \{1, 2, \dots, T\}$$
 (8)

其中 $\omega$ 表示叶子向量,q表示树的结构。正则项表示决策树的复杂度,复杂度定义为叶结点总数和叶节点权值平方和函数。

$$\Omega(f)_{t} = \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_{j}^{2} + \gamma T$$
(9)

其中, $\frac{1}{2}\lambda\sum_{j=1}^{T}\omega_{j}^{2}$ 表示 $\omega$ 的L2模平方, $\gamma T$ 表示叶结点总数。将式(8)和(9)代入目标函数(7)中,目标函数为

$$Obj^{(t)} \approx \sum_{i=1}^{n} \left[ g_{i} f_{t} \left( x_{i} \right) + \frac{1}{2} h_{i} f_{t}^{2} \left( x_{i} \right) \right] + \Omega \left( f_{t} \right)$$

$$= \sum_{i=1}^{n} \left[ g_{i} \omega_{q(x_{i})} + \frac{1}{2} h_{i} \omega_{q(x_{i})}^{2} \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} \omega_{j}^{2}$$

$$= \sum_{j=1}^{T} \left[ \left( \sum_{i \in I_{j}} g_{i} \right) \omega_{j} + \frac{1}{2} \left( \sum_{i \in I_{j}} h_{i} + \lambda \right) \varpi_{J}^{2} \right] + \gamma T$$

$$(10)$$

其中,I 被定义为每个叶子上面样本集合  $I_j = \{i | q(x_i) = j\}$ 。 定义  $G_j = \sum_{i \in I_j} g_i$ , $H_j = \sum_{i \in I_j} h_i$ ,从而

$$Obj^{(t)} \approx \sum_{i=1}^{n} \left[ g_{i} f_{t} \left( x_{i} \right) + \frac{1}{2} h_{i} f_{t}^{2} \left( x_{i} \right) \right] + \Omega \left( f_{t} \right)$$
$$= \sum_{i=1}^{T} \left[ G_{j} \omega_{j} + \frac{1}{2} \left( H_{j} + \lambda \right) \omega_{j}^{2} \right] + \lambda T$$

假设树的结构q固定,对 $\omega_i$ 求偏导,得到叶结点最优权重和最优的目标值

$$\boldsymbol{\varpi}_{j}^{*} = -\frac{G_{J}}{H_{J} + \lambda} \tag{11}$$

$$Obj = -\frac{1}{2} \sum_{j=1}^{T} \frac{G_J^2}{H_J + \lambda} + \gamma T \tag{12}$$

Obj 表示树的结构分数, Obj 越小,表示树的结构越好。对于当前的结点,采用贪心法进行子树的划分。每次尝试对已有叶结点加入一个分割点,枚举可行的分割点,选择目标函数最小、增益最大的划分。增益如式(13)

$$Gain = \frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{\left(G_L + G_R\right)^2}{H_L + H_R + \lambda} - \gamma \tag{13}$$

其中, $G_L$ 、 $G_R$ 分别表示左边树、右边树的增益之和。

# 3. 实验结果和讨论

#### 3.1. 电网负荷数据描述

电力系统负荷预报是电网系统建设、调度的重要依据,准确的电网负荷预报既能增强电力系统运行的安全性,又能改善电力系统运行的经济型。本文以电力系统负荷预报为应用研究背景,选取某市的历史小时负荷和温度观测数据[19],又根据历史负荷数据构造了前一周同一时刻负荷、前一天同一时刻负荷、前 24 小时平均负荷三个属性,即(1) prevWeekSameHourLoad =  $y_{i-1,j}$ : 前一周同一时刻负荷;(2) prevDaySameHourLoad =  $y_{i-1,j}$ : 前一天同一时刻负荷;(3) prev24HrAveLoad =  $\frac{1}{24}\sum_{j=1}^{24}y_{i,j}$ : 前 24 小时平均负荷。其中  $y_{i,j}$  表示第 i 天第 i 小时负荷(i 表示日期, i 表示小时)。完成属性构造以后,以干球温度、露点温度、一天中的小时、星期、是否为节假日、前一周同一时刻负荷、前一天同一时刻负荷、前 24 小时平均负荷作为输入变量,详见表 1,以第二天 24 点负荷作为输出,建立负荷预测模型。为了构建电网

Table 1. Description of power load data 表 1. 电网负荷数据描述

序号	输入变量	单位	均值	标准差	说明
1	干球温度	°C	50.04	18.66	
2	露点温度	°C	38.47	19.84	
3	时间	h			温度和负荷变量的采样时间序列 标记为1到24的整数
4	星期	day			标记为1到7的整数
5	是否为节假日				逻辑变量,如果该日期是节假日 标记为1,否则为0
6	前一周同一时刻负荷	MWh	15272.60	3010.66	
7	前一天同一时刻负荷	MWh	15270.29	3008.21	
8	前 24 小时平均负荷	MWh	15264.90	1775.04	

负荷预测模型, 共收集某地区电网负荷数据(2004~2008 年)样本 43847 条, 选取其中 35063 条数据 (2004~2007 年)作为训练样本数据建立负荷预测模型。剩余 8784 条负荷数据(2008 年)作为测试样本来检验所建模型性能。

# 3.2. 实验环境设置

电网负荷预报实验硬件平台为 2.59 GHz 的 Intel core i7 处理器,8G 内存,64 bit Win10 操作系统平台。模型实验编程语言为 python2.7,需要安装一些数学分析和建模、机器学习库,如 Numpy、Pandas、Matplotlib、XGBoost、Sklearn。本文数据分析和机器学习主要采用 XGBoost 工具,该软件包能够自动利用 CPU 的多线程并行计算,提升计算速度,同时引进模型复杂度控制和后期的剪枝处理改良算法准确度。

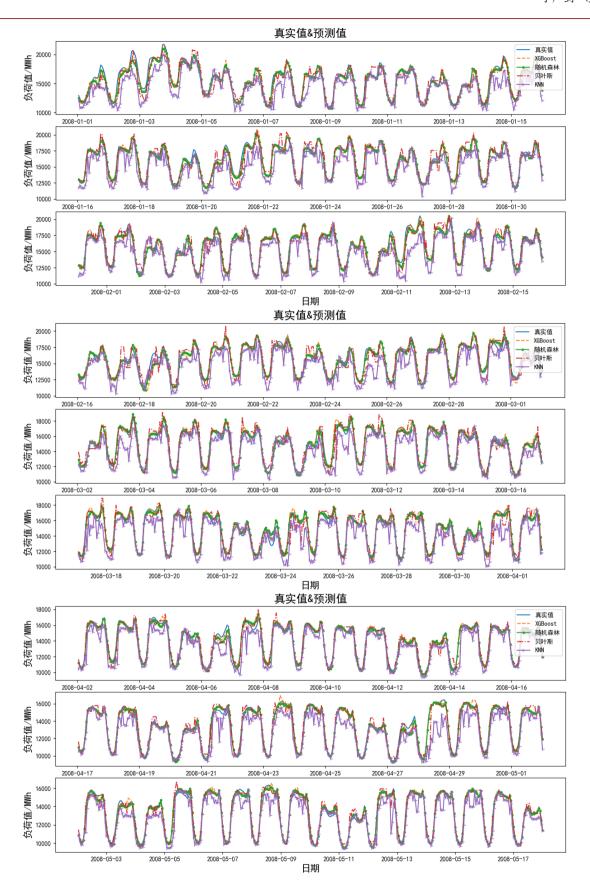
#### 3.3. 电网负荷预测模型构建

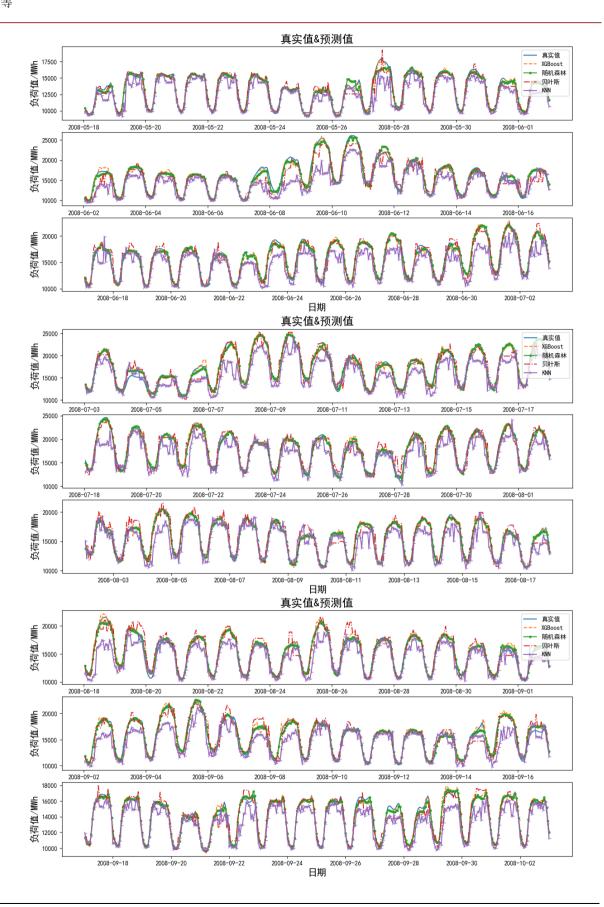
本文使用 XGBoost 工具包进行建模,首先加载电网负荷数据集。XGBoost 的参数可以分为三种类型:通用参数(General Parameters)、booster 参数以及学习目标参数(Task Parameters)。本文的电网负荷预报问题属于回归问题,构建的 XGBoost 模型目标函数参数选取 "count:poisson",即泊松回归的负对数似然,利用程序内置的交叉验证函数,划分 20%为验证集,80%为训练集,计算每次迭代的误差值确定最佳的num-round。通过交叉验证优化参数以后,最终确定模型的 booster 参数为: 树的最大深度(max-depth)为 6,迭代次数(num-round)为 150,收缩步长(eta)为 0.3。

#### 3.4. 实验结果和讨论

本文采用 XGBoost 算法、随机森林算法、贝叶斯算法和 KNN 算法预测了某市 2008 年 1 月 1 日到 2008 年 12 月 31 日连续一年的负荷,结果如图 1 所示。

好的预报模型,模型误差自相关函数曲线近似为一个白噪声序列。图 2 给出了不同建模方法得到的模型误差自相关函数曲线。为了进一步对比不同模型预测精度,将电网负荷的实际值与模型预测值分别作为横、纵坐标,画出四种方法预测值与实际值散点图,并求出各个散点图回归线的斜率,结果如图 3 所示。图 2 和图 3 实验结果均表明相比随机森林、贝叶斯和 KNN 方法,XGBoost 负荷预报模型性能最好。XGBoost、随机森林、贝叶斯以及 KNN 这几种建模方法得到的负荷预测值散点更加集中地分布在对角线





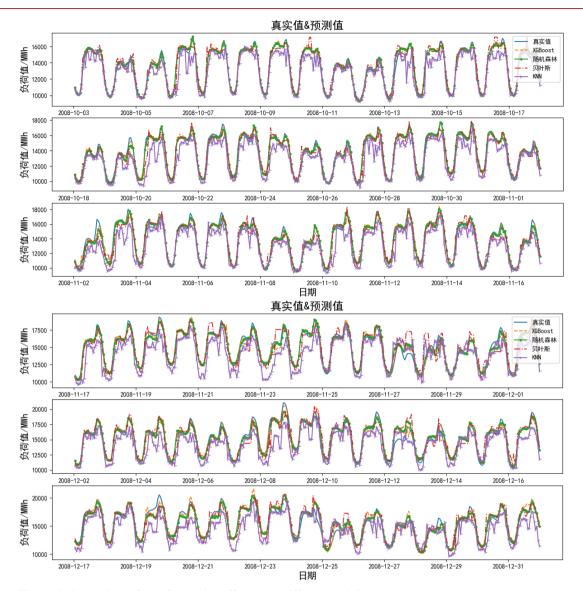


Figure 1. Comparison of load forecasting effect under different modeling method 图 1. 不同建模方法下负荷预测效果对比

附近,且回归线的斜率接近于 1,这进一步表明这四种模型得到的负荷值都比较接近于实际值。为了更好地描述不同算法模型的预测性能,采用平均绝对误差(MAE)、均方误差(MSE)、均方根误差(RMSE)、平均绝对百分误差(MAPE)四种评价指标作为各种方法预测效果判断的根据,即

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |E_i| = \frac{1}{n} \sum_{i=1}^{n} |Y_i - \hat{Y}_i|$$
 (14)

$$MSE = \frac{1}{n} \sum_{i=1}^{n} E_i^2 = \frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2$$
 (15)

RMSE = 
$$\sqrt{\frac{1}{n} \sum_{i=1}^{n} E_i^2} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2}$$
 (16)

MAPE = 
$$\frac{1}{n} \sum_{i=1}^{n} |E_i/Y_i| = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{(Y_i - \hat{Y}_i)}{Y_i} \right|$$
 (17)

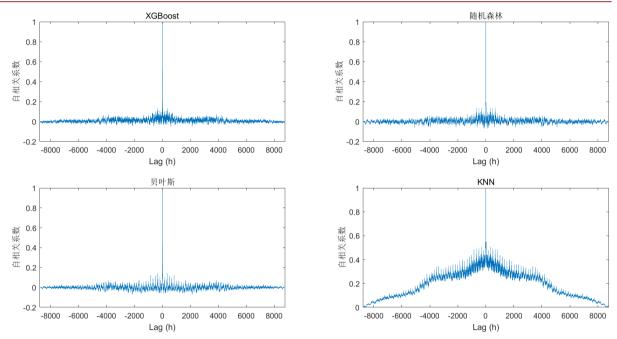


Figure 2. Autocorrelation function of estimating error of different models 图 2. 不同模型估计误差自相关函数

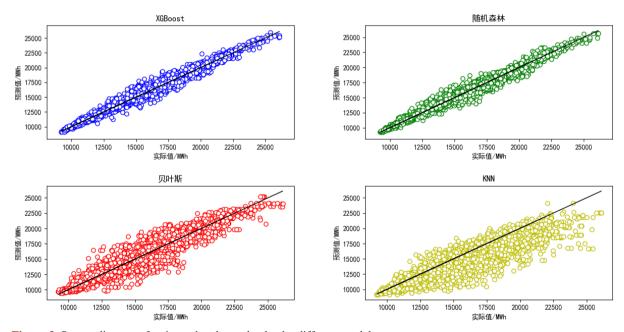
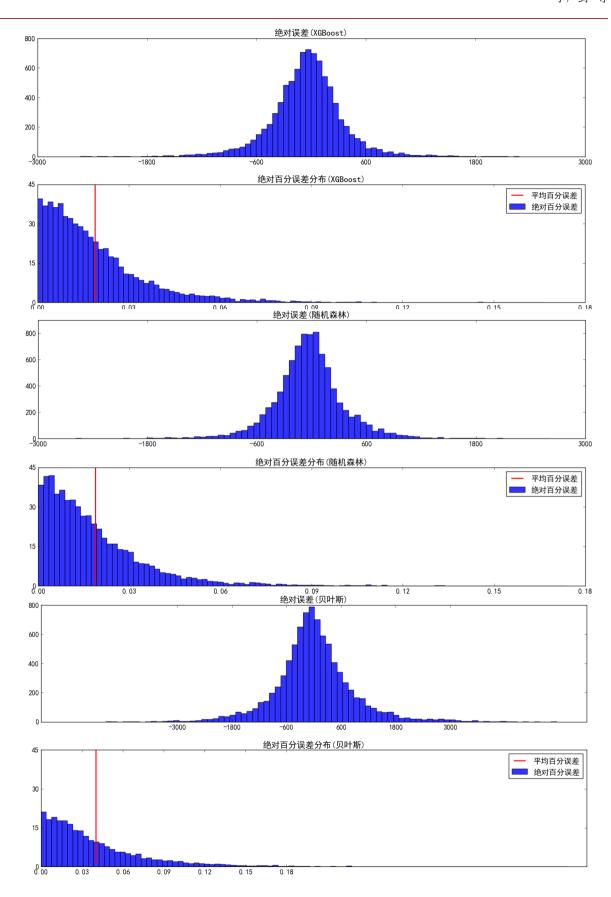


Figure 3. Scatter diagram of estimated and actual value by different models 图 3. 不同建模方法的估计值与实际值散点图

式中, $Y_i$ 表示第i个实际值, $\hat{Y}_i$ 表示第i个预测值, $E_i$ 表示第i个实际值与预测值的绝对误差。

图 4 给出的是四种算法模型预测结果的误差分布图。表 2 给出了四种模型下负荷预报误差和模型的训练时间结果。分析以上结果可见:

1) XGBoost 模型和随机森林模型得到的电网负荷的预测值相比其他两种方法能够更好地拟合真实值的变化趋势,且能保证较高的精度,具有更好的预测性能。XGBoost 和随机森林均采用集成学习框架,



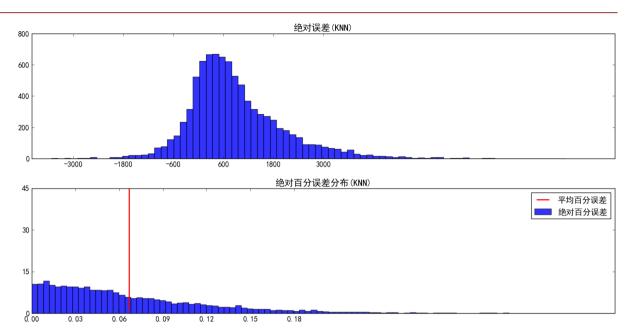


Figure 4. Predictive error distribution for different models

图 4. 不同模型预测误差分布

Table 2. Comparison of load forecasting error and model training time under different models 表 2. 不同模型下负荷预报误差和模型训练时间比较

模型名称	MAE	MSE	RMSE	MAPE	模型训练时间(s)
XGBoost	269.26	139,297.13	373.22	1.78%	3.03
随机森林	284.95	163,891.05	404.83	1.87%	16.96
贝叶斯	610.71	762,969.99	873.48	4.03%	1.57
KNN	1050.93	2,107,060.63	1451.57	6.67%	0.08

将弱分类器增强成强分类器。XGBoost 在迭代优化过程中使用了目标函数的二阶泰勒展开近似,以最快的速度收敛到全局最优,因此 XGBoost 精度更高。

- 2) XGBoost 和随机森林模型下得到的误差自相关函数曲线更加接近于白噪声序列。预测效果较理想。
- 3) 在预报误差方面 XGBoost 模型表现的最好,平均百分误差(MAPE)只有 1.78%,随机森林方法达到了 1.87%,贝叶斯和 KNN 算法误差较大。
- 4) 在模型的训练时间方面,KNN 模型表现的最好,只用了 0.08 秒,随机森林耗时长。XGBoost 原生语言为 C/C++,在进行节点的分裂时,支持各个特征多线程进行增益计算,因此算法计算速度更快。

综合来看,基于XGBoost模型的电力负荷预测结果在预测精度和计算时间综合指标方面优于其他方法。

#### 4. 结论

电力负荷预测是电力系统经济调度中的一项重要工作。本文以电力系统负荷预测为应用研究背景,采用 XGBoost 集成学习框架,利用其并行运算速度快、精度高、灵活性强、鲁棒性好的特点,构建了干球温度、露点温度、一天中的小时、星期、是否为节假日、前一周同一时刻负荷、前一天同一时刻负荷、前 24 小时平均负荷因子为输入未来一天 24 点电力负荷为输出的电网负荷预测模型。实验研究结果表明,该负荷预测模型预测结果能很好拟合实际负荷数据,平均百分误差在 2%以下。相较于贝叶斯、KNN 和随机森林方法,XGBoost 负荷预测模型具有预测精度高,模型训练时间短等优势,可在一定程度上满足

电网系统对电力负荷短期预报的工程需要,中长期电力负荷预报将在后续研究中逐步解决。

#### 参考文献 (References)

- [1] 李元诚, 方廷健, 于尔铿. 短期负荷预测的支持向量机方法研究[J]. 中国电机工程学报, 2003, 23(6): 55-59.
- [2] 廖旎焕, 胡智宏, 马莹莹. 电力系统短期负荷预测方法综述[J]. 电力系统保护与控制, 2011, 39(1): 147-152.
- [3] 彭鹏, 彭佳红. 基于多元线性回归模型的电力负荷预测研究[J]. 中国安全生产科学技术, 2011, 7(9): 158-161.
- [4] Ziel, F. and Liu, B. (2016) Lasso Estimation for GEF Com 2014 Probabilistic Electric Load Forecasting. *International Journal of Forecasting*, **32**, 1029-1037.
- [5] 金鑫, 李龙威, 季佳男. 基于大数据和优化神经网络短期电力负荷预测[J]. 通信学报, 2016, 37(Z1): 36-42.
- [6] Ceperic, E., Ceperic, V. and Baric, A. (2013) A Strategy for Short-Term Load Forecasting by Support Vector Regression Machines. IEEE Transactions on Power Systems, 28, 4356-4364.
- [7] Suganthi, L. and Samuel, A.A. (2012) Energy Models for Demand Forecasting: A Review. *Renewable and Sustainable Energy Reviews*, **16**, 1223-1240.
- [8] Ko, C.N. and Lee, C.M. (2013) Short-Term Load Forecasting Using SVR (Support Vector Regression)-Based Radial Basis Function Neural Network with Dual Extended Kalman Filter. *Energy*, **49**, 413-422.
- [9] 张素香, 赵丙镇, 王风雨. 海量数据下的电力负荷短期预测[J]. 中国电机工程学报, 2015, 35(1): 37-42.
- [10] 杜莉, 张建军. 神经网络在电力负荷预测中的应用研究[J]. 计算机仿真, 2011, 28(10): 297-300.
- [11] Mori, H. and Takahashi, A. (2011) Hybrid Intelligent Method of Relevant Vector Machine and Regression Tree for Probabilistic Load Forecasting. 2011 2nd IEEE PES International Conference and Exhibition on Innovative Smart Grid Technologies, Manchester, 5-7 December 2011, 1-8.
- [12] 雷正新, 韩蓓, 聂萌. 配电网大数据环境下的多点负荷预测算法与应用研究[J]. 电力系统保护与控制, 2016, 44(23): 68-78.
- [13] Chen, T., He, T. and Benesty, M. (2015) XGBoost: Extreme Gradient Boosting
- [14] Chen, T. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Sydney, 10-13 August 2015, 785-794.
- [15] 张昊, 纪宏超, 张红宇. XGBoost 算法在电子商务商品推荐中的应用[J]. 物联网技术, 2017, 7(2): 102-104.
- [16] 曹靖, 陈陆桑, 邱剑. 基于语义框架的电网缺陷文本挖掘技术及其应用[J]. 电网技术, 2017, 41(2): 637-643.
- [17] 樊鹏. 基于优化的 xgboost-LMT 模型的供应商信用评价研究[D]: [硕士学位论文]. 广州: 广东工业大学, 2016.
- [18] Friedman, J.H. (2001) Greedy Function Approximation: A Gradient Boosting Machine. Annals of Statistics, 29, 1189-1232.
- [19] Willingham, D. (2016) Electricity Load Forecasting for the Australian Market Case Study. https://www.mathworks.com/examples/matlabxl/community/19611-electricity-load-forecasting-australia



## 期刊投稿者将享受如下服务:

- 1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
- 2. 为您匹配最合适的期刊
- 3. 24 小时以内解答您的所有疑问
- 4. 友好的在线投稿界面
- 5. 专业的同行评审
- 6. 知网检索
- 7. 全网络覆盖式推广您的研究

投稿请点击: <a href="http://www.hanspub.org/Submission.aspx">http://www.hanspub.org/Submission.aspx</a>

期刊邮箱: sg@hanspub.org