

DOI:10.12158/j.2096-3203.2019.05.008

基于贝叶斯网络的电力客户停电敏感度预测

吕朋朋, 陶晓峰, 徐致光, 缪平, 熊霞, 毕善钰

(南瑞集团有限公司(国网电力科学研究院), 江苏 南京 211106)

摘要:精准预测停电敏感的电力客户群体,能够有效感知客户用电需求,提升客户用电满意度,助力提高电力服务水平。文中提出基于贝叶斯网络构建电力客户停电敏感度预测模型,从95598 客服平台、营销业务系统、用电信息采集系统获取分析数据,结合客户基本信息、用电信息、智能电能表计量信息以及用户用电交互行为,定义客户停电敏感度数据标签,对用户的停电投诉进行分析与预测。采用K折交叉验证法对停电敏感度预测模型进行实验验证。实验表明,基于贝叶斯网络构建的电力客户停电敏感度预测模型,在停电投诉分析应用中具备较高的精准度,验证了模型的有效性。

关键词:贝叶斯网络;停电敏感度;数据标签;K折交叉验证;后验概率

中图分类号:TM711;TP391.1

文献标志码:A

文章编号:2096-3203(2019)05-0049-06

0 引言

电力行业作为国民经济的基础产业,其稳定性不仅影响到公共安全与经济发展,也与广大人民群众的生活质量息息相关。当下,人民生活水平在不断提高,电力客户对供电可靠性的要求也越来越高,每一次的停电都可能激起客户的敏感神经,在供电服务过程中,稍有处理不慎,便可能引起敏感客户的停电投诉,甚至发展为舆情事件。随着智能电网建设的深入发展,电力公司信息化建设的逐步完善,这为分析电力客户的用电行为提供了有力的数据支撑。

在海量数据支撑的基础上,结合机器学习相关算法构建预测模型。文献[1—2]提出了著名的ID3、C4.5 决策树学习算法,以信息熵和信息增益度为衡量标准,实现数据的归纳分类;文献[3]总结了深度人工神经网络常用的算法,并对相关算法在数据分类与预测模型构建中的应用进行介绍;文献[4]提出了支持向量机算法,并证明了其在二分类任务中的卓越性能;文献[5]在支持向量机对二分类任务设计的基础上,推广了在多分类任务中的应用;贝叶斯决策论在机器学习、模式识别等诸多关注数据分析的领域都有极其重要的地位,文献[6—7]介绍了朴素贝叶斯分类器算法,并证明其在属性条件独立性假设的条件下取得相当好的性能;文献[8]证明了朴素贝叶斯分类器在信息检索领域有着较好的性能;文献[9]对贝叶斯分类学习过程以及

算法应用进行了详尽描述;文献[10]介绍了贝叶斯网络强大的表示能力和良好的可解释性,为不确定学习和推断提供了基本框架;文献[11]全面介绍了集成学习相关算法以及各类算法的实现过程及其应用。

文中结合贝叶斯分类器,结合电力客户基本信息、用电信息、电力交互信息以及智能电能表计量信息,定义电力客户停电敏感度数据标签,研究停电敏感预测模型,并基于K折交叉验证对预测模型进行了实验验证。停电敏感度的精准预测为电力公司服务部门实现管理创新、服务创新、效益创新提供了有力抓手。

1 基于贝叶斯网络的停电敏感度预测模型

文中以贝叶斯分类器相关理论知识为基础,研究与电力客户停电相关的电力数据项,定义停电敏感度数据标签,并构建停电敏感度分析模型。

1.1 贝叶斯分类相关理论

1.1.1 贝叶斯决策论

贝叶斯决策论是概率框架下实施决策的基本方法^[12]。对分类任务来说,在所有相关概率都已知的理想情形下,贝叶斯决策论考虑如何基于这些概率和误判损失来选择最优的类别标记。同时,贝叶斯方法不仅能够计算显式的假设概率,还能为理解多数其他方法提供一种有效的手段^[13—14]。

在给定 n 个样本 $x = (x_1, \dots, x_i, \dots, x_n)$,假设有 N 种可能的类别标记 $\{c_1, \dots, c_j, c_N\}$,期望基于样本集尽可能准确地估计出对每个样本 x_i ,使后验概率 $P(c_j|x_i)$ 最大的类别标记,结合贝叶斯定理, $P(c_j|x_i)$ 可写为:

收稿日期:2019-04-09;修回日期:2019-05-18

基金项目:国家电网有限公司科技项目“电、水、气、热能源计量一体化采集关键技术研究及应用”

$$P(c_j|x_i) = \frac{P(c_j)P(x_i|c_j)}{P(x_i)} \quad (1)$$

式中： $P(c_j)$ 为先验概率； $P(x_i|c_j)$ 为样本 x_i 相对于类标记 c_j 的类条件概率（或称为“似然”）； $P(x_i)$ 为用于归一化的“证据”因子（对所有类标记均相同）。因此，估计 $P(c_j|x_i)$ 的问题就转化为如何基于训练数据集来估计先验概率 $P(c_j)$ 和似然 $P(x_i|c_j)$ 的问题^[12]。

1.1.2 贝叶斯网络

基于式(1)来估计后验概率 $P(c_j|x_i)$ 时，类条件概率 $P(x_i|c_j)$ 是所有属性上的联合概率，这在现实有限的训练样本上是很难获取到的，朴素贝叶斯分类器采用了“属性条件独立性假设”^[15]，即，对已知类别，假设所有属性相互独立。半朴素贝叶斯分类器综合考虑现实任务中“属性条件独立性假设”成立可能性，适当考虑一部分属性间的相互依赖信息，从而既不需要进行完全联合概率计算，又不至于彻底忽略了比较强的属性依赖关系。贝叶斯网亦称“信念网”，借助于有向无环图来刻画属性之间的依赖关系，并使用条件概率表来描述属性的联合概率分布^[12]。根据对属性间依赖的涉及程度，朴素贝叶斯分类器不考虑属性间依赖性，贝叶斯网能够标识任意属性间的依赖性，而介于两者之间的则是一系列半朴素贝叶斯分类器。

贝叶斯网包括结构 G （有向无环图）和参数 Θ （属性间依赖关系）两部分，即，贝叶斯网 $B = \langle G, \Theta \rangle$ 。对于给定样本 $X = \{x_1, \dots, x_d\}$ ， x_i 为样本集 X_i 上的取值，假定 x_i 在 G 中的父节点集合为 χ_i ，则参数 Θ 描述了每个取值 x_i 的条件概率为：

$$\theta_{x_i|\chi_i} = P_B(x_i|\chi_i) \quad (2)$$

在贝叶斯网中，对于给定父节点集合，假设每个属性值与它的非后裔属性相互独立，于是 $B = \langle G, \Theta \rangle$ 可将属性值的联合概率分布定义为：

$$P_B(x_1, x_2, \dots, x_d) = \prod_{i=1}^d P_B(x_i|\chi_i) = \prod_{i=1}^d \theta_{x_i|\chi_i} \quad (3)$$

贝斯网络示例如图 1 所示，联合概率分布如公式(4)所示。

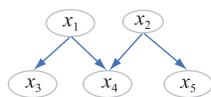


图 1 贝叶斯网络示例

Fig.1 Example of Bayesian network

$$P(x_1, x_2, x_3, x_4, x_5) = P(x_1)P(x_2)P(x_3|x_1)P(x_4|x_1, x_2)P(x_5|x_2) \quad (4)$$

1.2 停电敏感度分析模型构建

综合电力客户基本信息、电力特征信息以及相关行为数据，分析与电力客户的停电相关数据项，定义相关数据标签，并以此为基础构建停电敏感度分析模型。

1.2.1 电力客户停电敏感度数据标签

文中从属性、行为、评价 3 个方面定义电力客户停电敏感度标签。其中，属性主要包括基本属性、用电属性以及计量属性；行为主要包括用电行为以及交互行为；评价主要包括投诉评价以及停电敏感度的评价。电力客户停电敏感度数据标签如表 1 所示。

表 1 电力客户停电敏感度数据标签

Table 1 Power customer outage sensitivity label

| 标签分类 | 标签值 |
|------|--|
| 基本属性 | 姓名 年龄 性别 学历 家庭情况 从事行业 城市类别：城（可按经济水平细分）、镇、乡 用户类别：居民、工商业、专变、公变 |
| 用电属性 | 用电类别：重要备案、高耗能、预付费、电费担保等（多值标签） 供电电压：220 V, 380 V, 6 kV, 10 kV, 35 kV, 110 kV, 220 kV, 330 kV, 500 kV, 1 000 kV 负荷性质：高峰负荷（包括时间）、低谷负荷（包括时间）、平均负荷、负荷等级（从高到低分为 1—3 级别） 电价特征：分时电价、阶梯电价、合同电价等 用户（均）容量：针对公变总计户视为户均容量 立户年限：自用电日计 |
| 计量属性 | 计量设备信息：设备类型、接线方式、 计量方式、使用年限、更换情况、故障信息等 计量采集情况：电能示值、负荷曲线、 事件等业务应用数据采集情况 |
| 用电行为 | 用电趋势：某一时间段内用电走势、环比、同比信息 用电分布：尖、峰、平、谷、极谷用电情况 停电信息：停电类型、停电次数、停电时间、 停电时长、停电原因、停电通知方式等 |
| 交互行为 | 侧重于电力服务交互 交互方式：电话、营业厅 交互类别：起诉、投诉、建议、报修、需求 交互内容：电价、电费、有序用电、停电报修、服务态度等 交互态度：易激动、友好、一般 |
| 投诉评价 | 容忍度 敏感度 服务难度 投诉偏好 |

1.2.2 基于贝叶斯网络的停电敏感度预测模型构建

基于贝叶斯网络分析电力客户停电敏感度的过程,即确定贝叶斯网络结构 G 以及参数 Θ 的过程,文中结合电力系统中停电的影响因素,假定 G 的节点数一定(以上述六类数据为基础),在贝叶斯网络的学习过程中,在给定训练数据集 $D = \{x_1, x_2, \dots, x_d\}$ 的基础上,基于评分函数^[12](如式(5)所示)确定最优的分析电力客户停电敏感度的贝叶斯网,即最小化 $s(B|D)$ 。

$$s(B|D) = f(\theta) |B| - L(B|D) \quad (5)$$

$$L(B|D) = \sum_{i=1}^d \log P_B(x_i) \quad (6)$$

式中: $|B|$ 为贝叶斯网 B 包含的参数个数; $f(\theta)$ 为每个参数所需的字节数; $L(B|D)$ 为贝叶斯网 B 的对数似然。文中已知 $|B|$ 与 $f(\theta)$, 即,对于式(5)中第一项可视为常数(R),可将评优函数定义为式(7)所示。

$$s(B|D) = R - L(B|D) \quad (7)$$

$$\theta_{x_i|x_i} = \hat{P}_D(x_i|x_i) \quad (8)$$

此时,最优化贝叶斯网络等价于对参数 Θ 的极大似然估计,结合式(3),参数 $\theta_{x_i|x_i}$ 能够在训练数据集 D 上通过经验估计获得(见式(8)经验分布),从而确定分析停电敏感度的贝叶斯网络模型,具体过程如图2所示。

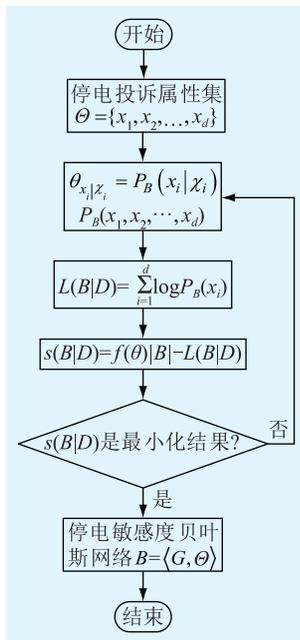


图2 基于停电敏感度的贝叶斯网络优化过程

Fig. 2 Bayesian network optimization process based on outage sensitivity

综上所述,基于样本训练集,结合式(6)一式

(8)得到分析客户停电敏感度的贝叶斯网,从而获取客户停电敏感度评价信息。

2 仿真实验

2.1 样本选择与评估

为了保证实验数据覆盖面尽可能广泛,充分考虑全国用电客户停电敏感情况,从全国各省 95598 工单数据随机抽取 10 000 工单信息(以每月数据为周期进行分析),关联营销业务系统获取用户基本信息,关联用电信息采集系统,获取用电信息,记为集合 D 。基于样本集合 D ,采用 K 折交叉验证法^[16]($K = 10$)划分训练集与测试集,即将样本集 D 随机划分为 10 个子集 $D = D_1 \cup D_2 \cup \dots \cup D_{10}$, $D_i \cap D_j = \emptyset (i \neq j)$, $|D_i| = |D_j| = 1 000$,每次选取其中一个子集作为测试集,其余 9 个子集作为训练集,依次选取不同的子集作为测试集,重复 10 次,最终将这 10 次测试结果的均值作为最终的验证结果。10 折交叉验证法如图 3 所示。

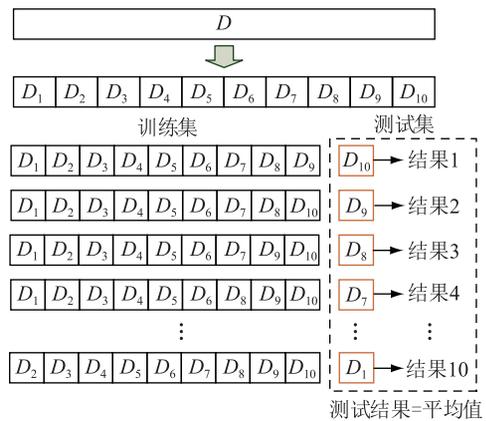


图3 10折交叉验证

Fig.3 10-Folding cross validation

2.2 仿真过程与结果验证

文中基于 Python 实现贝叶斯网络的构建,基于样本数据,得到用户停电投诉相关信息,分析停电投诉用户的基本信息、用电属性、计量属性以及用电行为,得到各属性节点的关联属性,形成初步的贝叶斯网络。结合式(7)和式(8)进行逐步优化,得到最优贝叶斯网络,同时,利用测试集进行校验,并给出测试验证结果与实际情况的对比情况。实验过程如下:

(1) 确定贝叶斯网络中各节点属性,记为 P ,并以停电敏感度属性作为输出属性;(2) 计算各属性节点在样本数据集中的概率值(占比);(3) 分别计算以各节点属性值为先验条件的条件概率,包括先验概率与后验概率值;(4) 对各节点后验概率设置阈值(P_λ),超过阈值则在对应节点之间连线,得到

初始贝叶斯网络 G_0 ; (5) 结合式(7)和式(8)对初始贝叶斯网络进行迭代优化; (6) 得到最终分析停电敏感度的贝叶斯网络 G 。

基于停电敏感度的贝叶斯网络构建实现过程首先需要初始化的标签信息包括停电投诉、用户类型、用电类型、城市类别、用电分布、电价特征、停电类型、停电时长、停电次数、交互方式、交互内容各属性值,分别记为 P_1, P_2, \dots, P_{11} , $P = \{P_1, P_2, \dots, P_{11}\}$, 设置后验概率阈值 $P_\lambda = 0.7$ 。关键代码逻辑如下:

- ① for P_i in P :
- ② P_i = 属性值对应样本在测试集中占比
- ③ for P_j in P :
- ④ 根据公式 1 计算 P_{ji}
- ⑤ if $P_{ji} > P_\lambda$:
- ⑥ 存在从属性 j 到属性 i 的边
- ⑦ end if
- ⑧ end for
- ⑨ end for
- ⑩ 得到初始的贝叶斯网络 G_0 。

输出:结合式(7)和式(8)迭代优化,得到停电敏感度的贝叶斯网络 G 。

根据上述构建过程,给出停电投诉属性节点的条件概率计算结果,如表 2 所示。其中,停电相关投诉率为 56.88%。

同样地,根据基于停电敏感度的贝叶斯网络构建过程,可以计算出各个属性值(停电投诉、用户类型、用电类型、城市类别、用电分布、电价特征、停电类型、停电时长、停电次数、交互方式、交互内容)条件概率分布^[17-20],并得到以这些属性为节点的贝叶斯网络,如图 4 所示。

在得到停电敏感度的贝叶斯网络后,采用 10 次 10 折交叉验证法进行验证,根据 10 次不同测试集进行验证,以 10 次测试结果的平均值做为最终的验证结果,基于贝叶斯网络预测停电投诉如图 5 所示,平均误差为 7.95%,整体测试准确率达到 92.05%。文献[21]提出基于 K -support 稀疏逻辑回归的停电敏感度预测模型预测结果准确率为 81.79%;文献[22]针对计划停电敏感用户进行分析,提出了基于机器学习技术的电力停电敏感客户标签体系,其分析准确率为 83.9%;文献[23-24]只针对频繁停电投诉研究了一种基于数据挖掘的频繁停电投诉预警模型,其预测准确率为 91.6%。综上所述,文中研究的基于贝叶斯网络的停电敏感度分析模型具备较好精度。

表 2 停电投诉条件概率
Table 2 Conditional probability of power outage complaints

| 属性 | 属性值 | 先验概率 | 后验概率 |
|-----------|-----------|-----------|-----------|
| 用户类型 | 居民 | 0.801 5 | 0.535 526 |
| | 非居民 | 0.199 5 | 0.763 118 |
| 用电类型 | 重要备案 | 0.010 2 | 0.951 108 |
| | 预付费 | 0.093 8 | 0.764 376 |
| | 电费担保 | 0.091 0 | 0.676 612 |
| 城市类别 | 普通 | 0.805 0 | 0.540 021 |
| | 城镇 | 0.527 2 | 0.588 791 |
| 用电分布 | 乡村 | 0.472 8 | 0.548 051 |
| | 峰用电 | 0.030 942 | 0.926 316 |
| 用电分布 | 平用电 | 0.874 473 | 0.539 304 |
| | 谷用电 | 0.057 314 | 0.555 366 |
| 电价特征 | 分时电价 | 0.881 153 | 0.527 912 |
| | 阶梯电价 | 0.157 349 | 0.971 770 |
| | 合同电价 | 0.034 634 | 0.907 834 |
| 停电类型 | 计划停电 | 0.276 547 | 0.280 442 |
| | 故障停电 | 0.723 453 | 0.937 144 |
| 停电时长 /min | ≥ 5 | 0.929 677 | 0.932 628 |
| | ≥ 15 | 0.809 951 | 0.955 611 |
| | ≥ 30 | 0.469 937 | 0.964 634 |
| | ≥ 60 | 0.170 007 | 0.993 834 |
| 停电次数 (月均) | ≥ 1 | 1 | 0.966 033 |
| | ≥ 2 | 0.881 681 | 0.961 649 |
| | ≥ 3 | 0.627 813 | 0.970 117 |
| | ≥ 4 | 0.209 916 | 0.991 694 |
| 交互方式 | 建议 | 0.000 130 | 0.003 455 |
| | 投诉 | 0.999 870 | 0.581 163 |
| 交互内容 | 电费 | 0.141 702 | 0.624 806 |
| | 服务态度 | 0.046 941 | 0.342 308 |
| | 人员违规 | 0.036 920 | 0.156 716 |
| | 业扩变更 | 0.006 329 | 0.150 000 |
| | 抢修质量 | 0.016 526 | 0.200 000 |
| | 接线错误 | 0.010 373 | 0.310 526 |
| | 电能质量 | 0.002 110 | 0.060 606 |
| | 营业厅服务 | 0.020 218 | 0.527 523 |
| | 电价 | 0.004 923 | 0.400 000 |
| | 其他 | 0.001 934 | 0.009 031 |

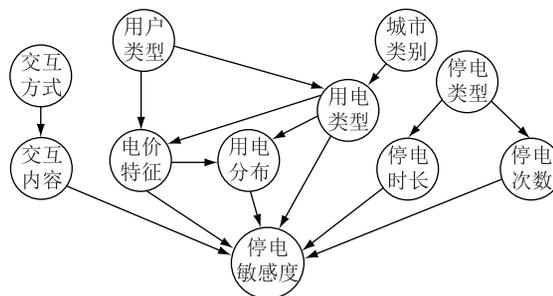


图 4 停电敏感度贝叶斯网络模型

Fig.4 Power outage sensitivity Bayesian network model

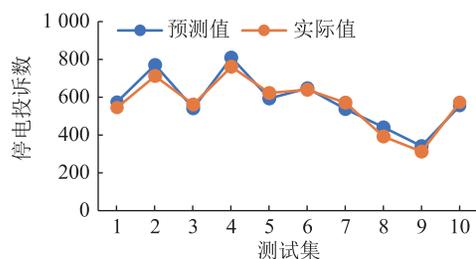


图5 基于贝叶斯网络预测停电投诉

Fig.5 Prediction of outage complaints based on Bayesian network

3 结语

文中定义了电力客户停电敏感度标签,以贝叶斯分类相关理论为基础,研究停电敏感度预测模型,从全国95598 客服平台、营销业务系统、用电信息采集系统中获取实验数据,对所构建的贝叶斯网络模型进行实验验证,实验表明,文中研究的贝叶斯网络模型在电力客户停电敏感度分析与预测中精准度较高。文中构建的停电敏感度预测模型的预测精准度在一定程度上依赖于后验概率阈值 P_λ ,该阈值的设定主要来源于经验数据,下一步需加强贝叶斯网络结构的学习能力在模型中的应用,能够做到自适应调整阈值。

参考文献:

- [1] QUINLAN J R. Induction of decision trees[J]. Machine Learning, 1986, 1(1): 81-106.
- [2] QUINLAN J R. C4.5: programs for machine learning[M]. Morgan Kaufmann, San Mateo, CA. 1993.
- [3] SCHMIDHUBER J. Deep learning in neural networks: an overview[J]. Neural Networks, 2015, 61: 85-117.
- [4] CORTES C, VAPNIK V. Support vector networks[J]. Machine Learning, 1995, 20(3): 273-297.
- [5] HSU C W, LIN C J. Errata to "a comparison of methods for multiclass support vector machines" [J]. IEEE Transactions on Neural Networks, 2002, 13(4): 1026.
- [6] DOMINGOS P, PAZZANI M. On the optimality of the simple-bayesian classifier under zero-one loss[J]. Machine Learning, 1997, 29(2-3): 103-130.
- [7] JORDAN A. On discriminative generative classifiers: a comparison of logistic regression and naive bayes[J]. Neural Processing Letters, 2008, 28(3): 169.
- [8] LEWIS DD. Naive (Bayes) at forty: the independence assumption in information retrieval[C]//In Proceedings of the 10th European Conference on Machine Learning (ECML), 1998, 4-15.
- [9] BISHOP C M. Pattern recognition and machine learning[M]. Springer, 2006.
- [10] PEARL J. Probabilistic reasoning in intelligent systems: networks of plausible inference[J]. Morgan Kaufmann, San Francisco, CA. 1988.
- [11] SCHWENKER F. Ensemble methods: foundations and algorithms [book review] [J]. IEEE Computational Intelligence Magazine, 2013, 8(1): 77-79.
- [12] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016. ZHOU Zhihua. Machine learning[M]. Beijing: Tsinghua University Press, 2016.
- [13] 付振罡,王明珠,王秀春. 电力行业全渠道客户体验测评研究[J]. 供用电, 2018, 35(1): 74-79. FU Zhengang, WANG Mingzhu, WANG Xiuchun. Research on whole channel customer experience evaluation in power industry[J]. Distribution & Utilization, 2018, 35(1): 74-79.
- [14] 马洲俊,樊飞龙,王勇,等. 基于多源异构数据的配电网故障信息挖掘与诊断[J]. 供用电, 2018, 35(8): 31-39. MA Zhoujun, FAN Feilong, WANG Yong, et al. Distribution network fault information mining and diagnosis based on multi-source heterogeneous data [J]. Distribution & Utilization, 2018, 35(8): 31-39.
- [15] TITTERINGTON D M, MURRAY G D, MURRAY L S, et al. Comparison of discrimination techniques applied to a complex data set of head injured patients[J]. Journal of the Royal Statistical Society, 1981, 144(2): 145-175.
- [16] DIETTERICH T G. Approximate statistical tests for comparing supervised classification learning algorithms[J]. Neural Computation, 1998, 10(7): 1895-1924.
- [17] 袁丹,王谊,李伟明,等. 基于分类模型的配电线路故障研判方法研究[J]. 浙江电力, 2018, 37(2): 11-15. YUAN Dan, WANG Yi, LI Weiming, et al. Research on fault diagnosis method for distribution line based on classification model[J]. Zhejiang Electric Power, 2018, 37(2): 11-15.
- [18] 陈伟华,王刚,叶仁杰,等. 基于复杂网络理论的系统脆弱度评估新方法[J]. 浙江电力, 2018, 37(3): 7-11. CHEN Weihua, WANG Gang, YE Renjie, et al. A new method for system vulnerability assessment based on complex network theory[J]. Zhejiang Electric Power, 2018, 37(3): 7-11.
- [19] 宁艺飞,陈星莺,谢俊,等. 公共楼宇群体价格型负荷响应的仿真分析方法研究[J]. 电力需求侧管理, 2019, 21(1): 42-46. NING Yifei, CHEN Xingying, XIE Jun, et al. Study on the method of simulation analysis for price based load response of public building[J]. Power Demand Side Management, 2019, 21(1): 42-46.
- [20] 盛银波,仲立军,张利庭,等. 基于停电明细数据的配电网可靠性监测与研究[J]. 浙江电力, 2017, 36(12): 70-74. SHENG Yinbo, ZHONG Lijun, ZHANG Liting, et al. Reliability monitoring and research of distribution networks base on detailed outage data[J]. Zhejiang Electric Power, 2017, 36(12): 70-74.
- [21] 耿俊成,张小斐,孙玉宝. 基于K-support稀疏逻辑回归的停电敏感度预测[J]. 计算机与现代化, 2018(4): 68-73. GENG Juncheng, ZHANG Xiaofei, SUN Yubao. Power failure sensitivity prediction algorithm using K-support sparse logistic regression[J]. Computer and Modernization, 2018(4): 68-73.

- [22] 蔡丽华. 基于机器学习技术的电力停电敏感客户标签体系[J]. 农村电气化, 2018(5):40-43.
CAI Lihua. Sensitive customer labeling system for power outage based on machine learning technology[J]. Rural Electrification, 2018(5):40-43.
- [23] 徐鑫,王莉,孙志杰. 一种基于数据挖掘的频繁停电投诉预警模型[J]. 信息记录材料, 2017,18(2):64-66.
XU Xin,WANG Li,SUN Zhijie. A frequent outage complaint early-warning model based on Data Mining[J]. Information Recording Materials, 2017,18(2):64-66.
- [24] 尹秀艳,张鸿,田新成,等. 基于大数据的台区重复停电监测分析[J]. 供用电,2017,34(1):38-41.
YIN Xiuyan,ZHANG Hong,TIAN Xincheng,et al. Monitoring

and analyzing of transformer districts with frequent power outage based on big data [J]. Distribution & Utilization,2017, 34(1):38-41.

作者简介:



吕朋朋

吕朋朋(1987),男,硕士,工程师,从事用电信息采集相关工作(E-mail:lvpengpeng@sgepri.sgcc.com.cn);

陶晓峰(1977),男,学士,高级工程师,从事电能计量及用电信息采集相关工作;

徐致光(1990),男,学士,工程师,从事用电信息采集相关工作。

Prediction of power customer outage sensitivity based on Bayesian network

LYU Pengpeng, TAO Xiaofeng, XU Zhiguang, MIU Ping, XIONG Xia, BI Shanyu

(NARI Group Co., Ltd.(State Grid Electric Power Research Institute), Nanjing 211106, China)

Abstract: Accurate prediction of sensitive power customer groups can perceive customer demand and improve customer satisfaction with electricity consumption and the level of power service effectively. A power customer outage sensitivity prediction model based on Bayesian network is proposed to predict the power customer outage complaints, which defines customer power outage sensitivity data labels by customer basic information, power consumption information, smart meter energy measurement information, and user power interaction behavior, coming from 95598 customer service platform, marketing system and power information collection system. It experimentally verifies the power outage sensitivity analysis model using K -folding cross validation method, shows that the power outage sensitivity prediction model based on bayesian network has high precision in the application of power outage complaint analysis, and the experimental results demonstrate the effectiveness of the prediction model.

Keywords: bayesian network; power outage sensitivity; data label; K -folding cross validation; posterior probability

(编辑 钱悦)