

基于大数据技术的电力营销策略分析研究

陈柘锜

(国网浙江诸暨市供电公司, 浙江 绍兴 311800)

摘要 在大数据技术快速发展的背景下, 电力市场的营销策略面临重要转型。本文探讨了大数据技术在电力营销中的应用, 重点分析了电力市场当前面临的挑战及大数据技术在营销策略分析中的优势。通过研究随机森林算法在电力客户群体划分中的应用, 展示了如何利用大数据优化营销策略, 提升客户满意度和企业竞争力。

关键词 大数据技术; 电力营销策略; 随机森林算法; 客户群体划分; 个性化营销

中图分类号: F425

文献标志码: A

文章编号: 2097-3365(2024)06-0073-03

随着科技的迅猛进步和数据量的急剧增加, 大数据技术已成为当前学术和商业领域的热门话题。在众多应用领域中, 电力行业特别是电力营销领域, 正逐渐成为大数据技术创新和应用的重要前沿^[1]。有效利用大数据技术对于电力企业来说, 不仅是提升竞争力的关键, 也是满足用户需求、优化资源配置的重要手段^[2]。因此, 深入分析基于大数据技术的电力营销策略, 对目前加快电力行业的发展有着重要的理论和实际意义。

本研究的核心动机在于探究大数据技术在电力营销策略分析中的作用, 基于大数据挖掘技术实现电力营销策略中的客户群体划分, 并基于此进行定制化的营销策略。旨在深入研究和探索大数据技术在电力营销领域的实际应用, 以提升电力企业的竞争力和服务质量。

1 大数据技术概述及其在电力营销中的应用分析

1.1 大数据技术内涵

大数据技术, 作为信息时代的产物, 涵盖了数据捕获、存储、分析、共享等多个方面, 用于处理那些传统数据处理应用软件难以应对的大规模数据集合。其主要意义在于从大量数据中能获取有价值的信息, 为决策制定提供强有力的支持。近年来, 随着计算能力的不断提升和算法的不断进步, 大数据技术在电力行业得到了广泛应用。尤其值得注意的是, 在电力营

销领域, 大数据技术的应用正逐步改变传统营销策略。

1.2 电力营销现状分析

当前的电力市场正经历着快速变革, 伴随着电力行业竞争发展和消费者需求的多元化, 电力营销面临着新的挑战 and 机遇。表 1 展示了电力营销领域的现状以及大数据技术对其潜在影响。

在大数据技术应用于电力营销的研究领域, 许多学者已经做出了重要贡献。这些研究主要集中在如何利用大数据改善客户关系管理、市场分析和风险控制等方面。例如, 一些研究专注于使用数据挖掘技术来分析用户行为, 从而更好地理解 and 预测用户需求^[3]。其他研究则探讨了如何通过分析大量的市场和运营数据来优化电力产品和服务^[4]。

1.3 大数据技术在营销策略分析中的应用优势

首先, 大数据技术可以帮助电力企业精确地划分客户群体。通过深度分析大规模的客户数据, 电力行业可以更准确地知道不同客户人群的特征和需求。通过对不同客户特征的挖掘和分析, 企业也可以将客户划分为不同的细分市场, 从而实现更精准的市场定位和目标客户群体的识别。其次, 大数据技术支持个性化营销策略。根据客户的特点和需求, 企业可以为不同客户提供个性化的产品和服务, 提高客户满意度和市场竞争力。最后, 大数据技术使得企业能够实时监

表 1 电力营销现状及大数据的潜在影响

电力营销要素	当前现状	大数据技术的潜在影响
市场定位	面向广泛的消费者群体, 缺乏精准定位	通过用户行为分析实现更精准的市场定位
产品定制	产品和服务相对单一	基于用户数据, 提供个性化的产品和服务
价格策略	价格相对固定, 缺乏弹性	利用市场数据, 实施动态和差异化定价
服务优化	服务方式传统, 用户反馈延迟	实时监控用户反馈, 快速优化服务响应

测用户反馈并快速改进服务。通过分析用户反馈数据，企业能够及时发现问题并采取提高服务质量，增强客户忠诚度。

2 基于大数据挖掘技术的用电客户群体划分方法

随机森林算法是一种基于集成学习的机器学习方法，它通过构建多个决策树并将它们的结果进行组合来进行分类和回归任务^[5]。在随机森林中，每个决策树都是通过随机采样数据集的一部分和特征的一部分来训练的，从而增加了模型的多样性。然后，通过对每个决策树的输出进行平均或投票来获得最终的预测结果。这种集成方法有助于减少过拟合风险，提高模型的稳定性和准确性。

电力营销策略的制定中，对于用电客户分群，随机森林算法可以将潜在的电力消费者划分为不同的群体，从而更好地理解他们的需求和行为模式。这有助于制定针对不同客户群体进行的定制化营销策略，以最大程度地提高销售效益。如下图1所示，为随机森林模型的构建过程：

首先，从原始数据集中随机抽取多个子样本。这个过程又被称为“Bootstrap 抽样”，其中每个子样本的大小通常与原始数据集相同。假设原始数据集表示为D，抽取出的子样本集为 D_1, D_2, \dots, D_n 。随后基于每一个子样本集构建决策树，在构建决策树时采用 CART 进行构建，CART 分类树算法使用基尼系数选择特征，基尼系数代表了模型的纯度，基尼系数越小，纯度越低，特征越好。对于某一个数据集D，其纯度用基尼值来度量：

$$Gini(D) = \sum_{i=1}^n p(x_i) * (1 - p(x_i))$$

$$= 1 - \sum_{i=1}^n p(x_i)^2 \quad (1)$$

其中， $p(x_i)$ 表示类别 x_i 出现的概率，n 是分类的

数目。Gini (D) 反映了从该数据集 D 中随机抽取两个样本，其类别标记不一致的概率。因此，基尼值越小，则数据集 D 的纯度越高。

随机森林中的每一棵决策树的构建过程如下所示：对于当前节点数据集 D，如果样本个数小于阈值或没有特征，则直接返回当前构建的决策子树。在计算得到各个特征（对应于不同类别用电客户的用电特点和情况）的基尼系数中，选择基尼系数最小的特征 A 作为最优的特征和划分点，并将父节点数据集划分为两部分，分别划分到对应的子节点处。对于树所产生的所有节点，重复以上决策树的构建过程，当每一个叶节点对应的数据集无法进一步划分时。完成 CART 决策树构建。

由于每一次进行决策树构建时利用的数据集都利用 Bootstrap 抽样获取的，因此可以充分利用构建的不同数据集，从海量数据中挖掘出用户用电需求和用户类型之间的潜在关系，融合不同的决策树的划分结果，随机森林模型的输出可以用于识别电力客户的关键行为特征，预测用电需求，从而实现精准营销。通过分析模型结果，电力公司可以识别出潜在的高价值客户，制定个性化的服务方案。

3 实验结果与电力营销策略分析

针对上述所采用的基于随机森林的大数据挖掘算法在电力营销方面的有效性和可靠性，本文通过随机森林算法对用电客户群体进行精准划分，并在此基础上探索基于该结果进行精准营销的可能性。研究对象为某电力公司的用户基本信息、历史用电记录和交易数据等。采用随机森林算法对客户数据进行分析。数据预处理包括数据清洗和归一化处理。特征选择侧重于影响客户用电行为的关键特征，包括用电量、总

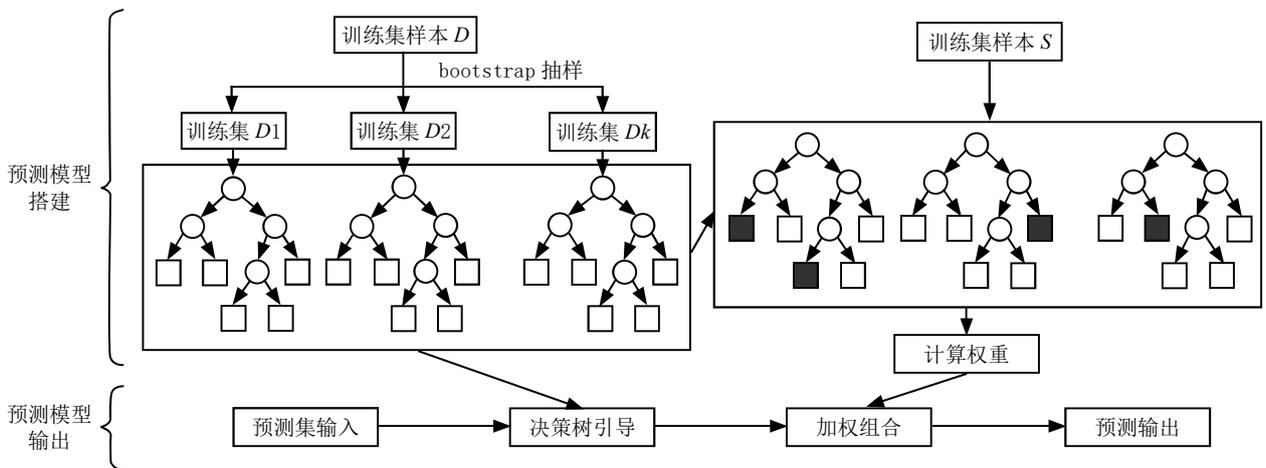


图1 随机森林模型结构

用电时间等。实验在高性能计算服务器上进行，使用 Python 语言和 Scikit-learn 库。其中为保证数据量纲对于实验结果的影响，采用 Z-score 进行标准化。

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

其中 μ 表示样本该特征的均值， σ 表示样本中该特征的标准差。

首先，从电力公司数据库中提取用户的客户数据，随后进行数据清洗，去除 NAN 等污染数据并处理异常缺失值。基于表 2 中获取的特征表示，并将其中的连续值进行离散化，以适应随机森林的输入，随后进行随机森林算法的训练，训练模型，并通过网格法对参数进行优化。

表 2 用于决策树构建的客户特征类别

特征类别	特征表示
用电行为特征	用电量（连续值）
	用电峰谷分布（连续值）
	用电时间模式（白天、夜晚、工作日、周末）
客户基本信息特征	用电历史
	地理位置（经纬度，海拔）
	家庭收入
	客户职业类型
	客户反馈历史

随后针对训练的模型，采用 5 折交叉验证的方法进行模型性能评估，采用准确率，召回率和 F1-分数作为评价指标，当准确率达 95% 以上，召回率达到 90% 以上时认为所构建的电力营销用户群体划分具有高度的准确性，采用 CART 决策树算法作为本文实验的基线模型，验证随机森林大数据挖掘算法是否可以有效提高客户群体划分结果的准确性，实验结果如表 3 所示。

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

表 3 不同数据挖掘算法的客户群体划分结果

数据挖掘算法	准确率	召回率	F1-score
CART 决策树	90.2%	89.9%	0.88
随机森林	96.3%	93.5%	0.91

上述客户群体划分中识别目标为两个主要的客户

群体，高用电量企业客户和低用电量家庭用户，因此在两种数据挖掘算法中准确率都有较为良好的表现，对于随机森林算法，经过网格法的参数调优，可以实现 95% 以上的客户群体划分准确率，以及 93.5% 的不同客户群体召回率，相对于基线模型提升效果显著，表明随机森林大数据挖掘算法可以有效提高客户群体划分结果的准确性。

对于高用电量企业客户，随机森林模型表明其通常在工作日的白天用电负载大。基于这一发现，电力公司可以为这些客户提供优惠的工作日白天电价，以及相关的能效改进咨询服务。对于低用电量家庭用户，模型显示他们的用电量在夜间和周末较高。针对这一群体，电力公司可以推出夜间和周末的优惠电价计划，以及家庭用电安全和节能的教育活动。利用随机森林模型的结果，电力公司能够更精确地了解不同客户群体的用电习惯和需求，从而制定更加个性化和有效的营销策略。有望提高客户满意度，优化电力资源的分配和利用。

4 结束语

本文通过深入分析大数据技术在电力营销策略中的应用，探讨了大数据技术如何优化电力市场的客户群体划分和营销策略。研究显示，利用大数据技术，尤其是随机森林算法，能有效提高客户群体划分的准确性，从而支持电力企业制定更加个性化和有效的营销策略。这不仅可以提高企业的市场竞争力，还可以优化资源配置，增强客户满意度。

参考文献:

- [1] 殷悦. 大数据在电力营销中的应用 [J]. 科学与信息化, 2023(16):13-15.
- [2] 卢平. 基于大数据技术的电力营销系统应用研究 [D]. 武汉: 华中科技大学, 2020.
- [3] 赵娟. 基于大数据的电力营销信息化技术应用 [J]. 电力系统装备, 2023(06):170-172.
- [4] 武新娟, 薛建德, 迪里达尔·库尔班, 等. 电力大数据在供电客户精准营销中的应用 [J]. 中文科技期刊数据库 (全文版) 经济管理, 2023(07):4-6.
- [5] 钟楚彦. 基于随机森林模型的精准营销研究 [J]. 经济管理文摘, 2020(03):169-170.