

[电力营销]

# 基于灰色关联分析的火力发电企业电力营销信息多标签分类研究

文/钟懿敏

广东惠州平海发电厂有限公司

由于电力营销信息的变化频繁且数据维度高，难以捕捉到稳定的模式，其复杂性和多样性增加了分类难度，导致在进行分类时ROC曲线表现较差。因此，研究提出基于灰色关联分析的火力发电企业电力营销信息多标签分类方法。首先，确定包含多个关键维度的电力营销信息参考数列，并对参数进行标准化处理。其次，在因电力营销信息变化频繁而导致数据样本有限的情况下，利用灰色关联分析方法对于小样本数据具有较好的适应性的特点，计算各企业电力营销信息与参考数列之间的关联程度。最后，基于计算结果，设定分类阈值并分配特定的分类标签，采用动态调整方法优化分类阈值，实现对火力发电企业电力营销信息的多标签分类。实验结果表明，该方法能够有效处理电力营销数据，ROC曲线更加贴近理想状态，说明该方法具有更高的分类准确度和更强的泛化能力。

电力营销是指电力企业利用市场机制组织商品电力的生产、购销和分配，推动电力交易，实现市场经济和市场定价的经济活动。随着市场竞争的加剧，精准了解和分析客户需求成为电力企业制定有效营销策略的关键。客户分类作为营销管理的核心环节，对提高营销效率和客户满意度至关重要。

现阶段，已有相关学者设计了一系列的电力信息标签分类方法。例如文献[1]中通过对客户的用电行为、消费习惯和信用记录等多维度信息进行综合分析，利用数据挖掘和机器学习技术构建客户分类模型，将客户群体细分为优质客户群、待发展客户群及基础客户群，以提供个性化服务和营销策略；文献[2]中设计基于量子竞争决策算法的电力营销信息多标签分类方法，利用量子计算的并行性和高效性，快速处理和分析电力营销信息，构建多标签分类模型，不仅考虑客户用电行为，还结合市场趋势和政策变化等外部因素，使分类结果更具前瞻性和实用性。然而，上述方法在构建分类模型时主要依

赖历史数据和经验知识，缺乏对市场动态变化的实时感知和适应能力。

针对传统研究的不足，本文提出基于灰色关联分析的火力发电企业电力营销信息多标签分类研究方法，综合考虑客户用电行为、市场趋势和政策变化等多种因素，通过计算各因素之间的关联度，实现对客户信息的多维度标签分类。

## 确定电力营销信息参考数列

由于火力发电企业的电力营销信息复杂多样，包含多个维度和标签，为建立一个统一且可比较的基准，首先确定电力营销信息的参考数列。这些参考数列涵盖了发电量、售电量、电价结构、市场需求变化、客户满意度、政策导向及竞争态势等多个关键维度，全面反映了企业的核心竞争力与市场定位。为确保参考数列的代表性和全面性，对这些维度进行深入梳理，并依据当前电力市场的特性和企业营销战略的重点，从原始数据池中挑选出对营销

效果起主导作用的参数作为参考数列，见表1。

表1 电力营销信息参考数列

| 参考数列名称    | 描述                           |
|-----------|------------------------------|
| 月度售电量增长率  | 反映企业营销活动成效与趋势的核心指标           |
| 发电量增长率    | 体现企业生产能力的关键指标，尤其在电力需求高峰期     |
| 燃料成本      | 影响企业盈利能力和电价结构的重要因素           |
| 发电设备利用率   | 反映企业设备使用效率和生产效率的指标           |
| 客户满意度指数   | 衡量客户对企业服务满意度的量化指标            |
| 客户构成比例    | 各类客户的占比情况                    |
| 电价结构合理性   | 评估企业电价设置是否合理，影响客户选择 and 市场份额 |
| 可再生能源发电占比 | 反映企业绿色转型和可持续发展能力的指标          |
| 电力市场交易活跃度 | 衡量企业在电力市场交易中的活跃程度和参与度        |
| 电力输送损耗率   | 体现企业电网运营效率和成本控制能力的指标         |
| 电力服务质量指标  | 包括供电可靠性、停电恢复时间，影响客户满意度和忠诚度   |
| 市场营销费用投入  | 企业为提升市场份额和品牌知名度而投入的费用        |

为确保参考序列与比较序列之间的可比性，对所有选定的参数进行标准化处理。这一过程可以表示为

$$Z_{ij} = \frac{Z'_{ij} - Z_{\min}}{Z_{\max} - Z_{\min}} \quad (1)$$

式中， $Z_{ij}$ 为标准化后的数据； $Z'_{ij}$ 为原始数据池中的各个数据点，其中 $i$ 表示不同的企业， $j$ 表示不同的营销指标； $Z_{\max}$ 、 $Z_{\min}$ 分别为对应的最大、最小值。

通过这种处理方式，能够得到一个相对值序列，有助于消除单位与数量级差异对分析结果的影响，使得各因素之间的比较更加公平、准确。

### 计算各数列灰色关联程度

为量化不同电力营销信息数列与参考数列之间的关联程度，本文采用灰色关联分析的方法。假设关联系数 $\zeta_{ij}$ 表示第 $i$ 个企业的第 $j$ 个指标与参考数列的关联程度，该关联系数可以通过下式展开计算

$$\zeta_{ij} = \frac{\rho \max_{ij} |Y_i - Z_{ij}| + \min_{ij} |Y_i - Z_{ij}|}{\rho \max_{ij} |Y_i - Z_{ij}| + |Y_i - Z_{ij}|} \quad (2)$$

式中， $\rho$ 为分辨系数； $Y_i$ 为火力发电企业电力营销信息的参考数列； $\max_{ij} |Y_i - Z_{ij}|$ 和 $\min_{ij} |Y_i - Z_{ij}|$ 分别为所有

比较序列与参考序列之间差值的最大值和最小值<sup>[3]</sup>。

为全面评估各企业在电力营销信息上与参考数列的整体相似程度，进一步计算关联度 $r_i$ 。关联度 $r_i$ 表示第 $i$ 个企业与参考数列的整体关联程度，计算公式为

$$r_i = \frac{1}{m} \sum_{j=1}^m \omega_j \zeta_{ij} \quad (3)$$

式中， $m$ 为指标的数量； $\omega_j$ 为第 $j$ 个指标的权重。

通过计算各数列的灰色关联程度，能够精确量化不同企业的电力营销信息与参考数列之间的相似性和差异性。

### 输出电力营销信息多标签分类结果

根据灰色关联度的计算结果，本文设定合理的分类阈值，将关联度划分为不同的区间，进而为每个区间分配特定的分类标签。这一流程如图1所示。

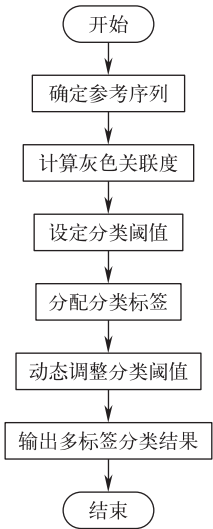


图1 火力发电企业电力营销信息多标签分类流程

为确保分类的准确性和有效性，基于深入的业务需求分析和详尽的数据分析结果，设定 $k$ 个分类阈值 $\theta_1, \theta_2, \dots, \theta_k$ ，这些阈值将关联度划分为 $k+1$ 个区间： $[\theta_0, \theta_1), [\theta_1, \theta_2), \dots, [\theta_k, \theta_n]$ 。为准确反映火力发电企业在电力营销信息方面的特征和表现，为每个区间 $[\theta_{j-1}, \theta_j)$ 分配了一个具体的分类标签 $L_j$ ，其中 $j=1, 2, \dots, k+1$ 。

考虑到不同企业在营销信息上的差异性，进一步采用动态调整分类阈值的优化方法。设调整因子向量为 $A=(A_1, A_2, \dots, A_n)$ ，其中 $A_i$ 表示第 $i$ 个企业的调整因子。优化调整后的分类阈值 $\theta'_j$ 计算公式为

$$\theta_j' = \theta_j + \sum_{i=1}^n A_i \omega_{ij} \quad (4)$$

式中,  $\omega_{ij}$  为第  $i$  个企业在确定第  $j$  个分类阈值时的权重。

这些权重可以根据火力发电企业的特定属性进行设定, 从而确保调整后的分类阈值更加符合企业的实际情况。最终, 多标签分类结果可以表示为

$$\text{Label}_i = \{L_j | \theta_{j-1}' \leq r_i < \theta_j', j=1, 2, \dots, k+1\} \quad (5)$$

式中,  $\text{Label}_i$  表示第  $i$  个火力发电企业的电力营销信息分类标签集合。通过这一过程, 能够精确地将各火力发电企业的电力营销信息划分为多个标签, 为企业提供更加全面、细致的营销信息分类服务。由此, 完成了本文基于灰色关联分析的火力发电企业电力营销信息多标签分类的研究。

## 实验与结果分析

### 1. 实验准备

本实验旨在探索火力发电企业电力营销信息的多标签分类方法。实验过程中, 由于要处理大量数据和进行模型训练, 所以需要确保计算机硬件资源足够强大且稳定, 以防止实验因资源短缺或设备故障而中断, 或导致结果不准确。为此, 本实验采用高性能的计算机/服务器, 配备 Intel Xeon 处理器、256 GB ECC 内存、快速存储设备和千兆网络接口, 同时还安装了 Ubuntu 操作系统、Python 编程环境和 MySQL 数据库, 以确保实验能够高效运行且数据安全可靠。同时, 考虑到火力发电企业的营销数据包含许多敏感信息和商业秘密, 所以在数据的收集和处理阶段, 必须严格遵守数据保密协议, 确保数据的安全性和隐私得到充分保护。

### 2. 实验样本数据集

本文实验环境基于某电力企业的实际运营背景, 精心挑选了该企业运营过程中产生的多组电力营销数据集作为实验对象。这些数据集具体信息见表 2。

表 2 中, 数据集编号用于唯一标识每个数据集, 样本数表示各数据集中包含的客户记录总数, 涵盖客户在电力消费、缴费行为和用电习惯等多个维度的信息。类别数 (标签数) 揭示了各数据集中用于描述客户特征的标签数量。这些标签是多标签分类任务的核心, 能够反映客户的不同需求和偏好。这些数据集不仅为实验提供丰富的数据资源, 还确保

实验结果的可靠性和准确性。本文将在后续的实验用于训练和测试多标签分类模型, 以探索最优的分类方法和策略。

表 2 实验样本数据集详情

| 数据集编号 | 样本数    | 类别数 (标签数) |
|-------|--------|-----------|
| DS1   | 20 000 | 6         |
| DS2   | 18 000 | 5         |
| DS3   | 22 000 | 11        |
| DS4   | 19 500 | 5         |
| DS5   | 21 000 | 9         |
| DS6   | 17 800 | 3         |
| DS7   | 23 000 | 12        |
| DS8   | 20 500 | 8         |

### 3. 实验结果及分析

为验证研究方法的优越性, 将本文方法设为方法 1, 将现有的电力营销客户标签分类方法优化设计设为对比方法 2, 同时, 引入基于量子竞争决策算法的电力营销信息多标签分类方法作为方法 3 进行对比。在本次实验中, 选用 ROC 曲线作为关键的评价指标, 以全面评估三种方法的性能。三种方法在 ROC 曲线上的具体表现如图 2 所示。

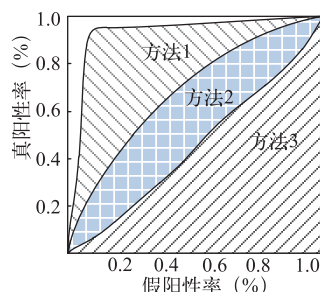


图 2 三种方法在 ROC 曲线上的表现

观察图 2 可以看出, 三种方法在 ROC 曲线上的表现存在显著差异, 本文方法在 ROC 曲线上的表现尤为突出, 显著优于其他两种方法。具体而言, 方法 3 的 ROC 曲线位置相对较低, 这直观地反映出其分类性能相对较弱。尽管在低假阳性率区域, 方法 3 的真阳性率有所上升, 但整体而言, 其分类准确性仍有较大的提升空间。相比之下, 方法 2 的表现虽然优于方法 3, 其 ROC 曲线位置相对较高, 但在真阳性率与假阳性率的平衡上仍有提升空间。而本文方法的 ROC 曲线在坐标轴中占据了最高位置, 且紧贴左上角, 这充分展示了其卓越的分类性能。

(下转第 45 页)

效措施。在风电机组的安装过程中，可以使用减震支架、减震垫等装置来减少机组的振动。这些装置通过其弹性材料和结构设计，能够有效降低机组的振动传递到周围环境，还可以在机组周围增加吸音材料来吸收噪声并降低振动对周围环境的影响。

(3) 运行状态的调整

转速太低或太高都可能导致振动现象的发生。因此，应定期检查转速控制系统的运行情况，及时发现问题并进行修理。同时，要根据实际情况对转速进行调整，使其尽量符合设计要求，以减少振动问题的发生。

结束语

本文首先通过受力分析对风电机组振动产生的原因和振动故障的主要类型；其次分别对时域分析和频域分析两种振动故障检测技术的具体应用进行详细介绍，重点在算法的分析和应用效果；最后，结合常见的振动故障和风电机组的结构特点，从主要部件的维护、稳定结构的加固和运行状态的调整

三个方面给出可行的治理对策。**EA**

(收稿日期：2024-07-01)

参考文献

[1] 何玉灵, 徐明星, 张文, 等. 气隙静偏心故障对永磁风力发电机定子绕组振动的影响[J]. 中国工程机械学报, 2024, 22 (1): 83-88.

[2] 邓亦峰. 风电场双馈风力发电机异常振动故障及排除措施[J]. 仪器仪表用户, 2023, 30 (10): 95-97.

[3] 朱广贺, 朱智强, 袁逸萍. 基于振动传感器的风力发电机故障检测算法[J]. 传感技术学报, 2023, 36 (1): 108-112.

[4] 张利慧, 李晓波, 俎海东, 等. 风力发电机组滚动轴承故障振动诊断方法综述[J]. 内蒙古电力技术, 2022, 40 (3): 7-12.

[5] 魏协奔, 卢旭锦, 孙培明, 等. 浅析风力发电机组振动状态监测与故障诊断[J]. 中国设备工程, 2021, (15): 142-143.

(上接第 41 页)

在低假阳性率的情况下，本文方法依然能够保持较高的真阳性率，充分证明了其在捕捉客户标签关键信息和处理复杂电力营销数据方面的显著优势。这种优势源于本文方法在特征提取和分类算法上的设计。本文方法通过灰色关联分析的方式更准确地提取特征和优化分类算法，灰色关联分析可以根据已有的少量数据来分析不同用户用电行为之间的关联程度，且其能够动态适应数据的变化。这使得本文方法能够更精确地识别客户标签，从而在实际应用中展现出更高的实用价值。

结束语

本文研究构建电力营销信息数据集，采用灰色关联分析方法计算指标间关联度，成功实现客户信息的多维度标签分类。该方法有效提升分类准确性和稳定性，为电力企业描绘出深入且全面的客户画像，助力企业实施更有针对性的营销计划与服务优化措施。通过这一方法，企业能更好地理解客户需求，优化服务体验。未来，将继续探索更先进的灰色关联度分析技术和计算模型，力求更深层次地

挖掘数据潜力，增强分类的稳定性和可靠性，为电力企业的营销决策提供有力支持，推动企业持续发展。**EA**

(收稿日期：2024-11-21)

参考文献

[1] 梁哲辉, 殷秀颜, 李韞莛. 电力营销客户标签分类方法优化设计[J]. 信息技术, 2023(5): 148-152.

[2] 杨峰, 罗尉丹, 李仰杰. 基于量子竞争决策算法的电力营销信息多标签分类方法[J]. 微型电脑应用, 2023, 39(3): 100-103.

[3] 曹斌. 基于决策树分类算法的电力营销数据异常识别方法[J]. 长江信息通信, 2024, 37(9): 99-101.

[4] 赵天宇. 电力市场化改革对新能源发电企业的影响分析及策略[J]. 中国战略新兴产业, 2024(23): 103-105.

[5] 江腾飞. “新电改”背景下YF电厂电力营销策略探究[J]. 能源科技, 2024, 22(2): 90-93.