

# 聚类分析在电力客户分类中的应用分析

王泽楠 李新华

内蒙古电力(集团)有限责任公司鄂尔多斯供电分公司 内蒙古自治区鄂尔多斯市 017000

DOI: 10.12238/ems.v7i8.14711

**[摘要]** 随着电力工业的发展,供电公司的用电客户越来越多且需求越来越复杂多样;而客户服务中心作为与用户接触最直接的部门之一,则是供电企业的重要组成部分,在保证售电收入和优质服务的同时为客户提供更好的产品和服务显得尤为重要。通过对电力客户的分类研究可以更合理地利用现有资源来满足不同层次用户的需要,从而提高企业的市场占有率。本文详细阐述了聚类分析的基本原理、常用算法及其在电力客户分类中的具体应用,同时分析了应用过程中存在的问题及解决对策,并对未来发展趋势进行了展望。

**[关键词]** 聚类分析; 电力客户分类; 数据挖掘; 应用分析

## 一、引言

随着电力市场环境的变化及竞争压力加大,如何对客户细分并为其提供优质个性化服务成为当前迫切需求。通过电力用户分类研究可以了解不同类型用户的用电特点以及其潜在的需求与偏好,并根据不同的类别制定相应的营销策略和服务方案来提升用户满意度和忠诚度。因此,建立一套科学合理的电力用户分类模型对于促进电力建设发展有着十分重要的意义。聚类分析作为一种无监督学习方法,能够在大量的电力客户数据中自动发现数据的内在结构和规律,为电力客户分类提供了有效的技术支持。

## 二、常用聚类算法及其特点

### 1. K-Means 算法

K-means 算法是最经典的划分型聚类算法,在电力客户分类中的应用也比较广泛。K-means 算法的主要思路为:先从样本集合  $S$  中任意选取  $k$  个元素作为初值聚类中心,再把所有数据点按与各聚类中心的距离划分为  $k$  类;之后根据各簇内的所有数据对象重新计算出新的聚类中心,并将其替换原来的聚类中心继续执行上述步骤,直到相邻两次迭代后的聚类中心变化小于某一阈值或者达到了预先设定的最大迭代次数为止<sup>[1]</sup>。K-means 算法的优点在于算法简单易实现且计算速度较快,对于大容量的数据集有很好的适应能力。然而,该算法也存在一些缺点,例如对初始聚类中心的选择较为敏感,不同的初始聚类中心可能导致不同的聚类结果;需要事先指定聚类的个数  $K$ ,而  $K$  值的确定往往比较困难,若  $K$  值

选择不当,会影响聚类效果。

### 2. 层次聚类算法

根据距离定义的不同,层次聚类又可以划分为凝聚式层次聚类 (AGNES)、分裂式层次聚类 (DIANA) 以及它们的混合形式 BIRCH, CLARA 等。层次聚类的基本思想是将样本按照一定的规则进行分组,并构造出由不同粒度的多个子簇构成的多层次的分类体系,包括两方面的内容:一是如何选择两个近邻样本作为一类;二是如何确定某一时刻应该有几类。由于这种划分方法无需预先知道类别数目并能产生较好的层次结构图,因此被广泛应用于实际问题的研究当中。但是该算法存在一些不足之处:对于大规模的数据集来说,的运行时间比较长;当聚类完成之后就无法改变,可能导致聚类结果不理想<sup>[2]</sup>。

### 3. DBSCAN 算法

DBSCAN

(Density-Based Spatial Clustering of Applications with Noise) 算法属于一种基于密度的聚类方法。它把数据的空间划分成核心对象、边沿对象以及噪声三个部分:核心对象是指在它的邻域中包含了足够多的对象;而边沿对象则是指在其邻域中的对象数目达不到要求,但是却处于一个已知的核心对象的邻域内的对象;最后就是噪声对象,即那些不属于前两类的对象。DBSCAN 算法是从每个核心对象出发逐渐向外扩张从而产生簇,可以找出任何形状的簇并且可以区分出噪声点 (DBSCAN 算法根据是否为簇成员判断是否为噪声),因此

适合于含有大量噪点的数据集,并能自动确定簇的数量。缺点在于需要人为设定一些超参量(如邻域半径和最小点数),这些超参量选择不当将会严重影响到聚类的效果。

#### 4. 其他聚类算法

除了以上提到的一些常用聚类方法之外,还有其他一些聚类方法也常用于电力客户的分类研究,如基于网格的聚类方法,首先把整个数据空间划分成若干个矩形格子(即网格),然后通过处理这些网格而得到最终的结果;该方法的优点是计算速度较快且对于输入的数据序列没有要求;再比如模糊C均值聚类方法,它是传统K均值的一种改进形式,其特点是每个样本可以同时被分配到多个类别中,并赋予相应的权重或隶属度,这样就使得数据的不确定性和模糊性得到了有效的表达与描述<sup>[3]</sup>。每种聚类方法都有自己的特点及优势,在实际应用中需要根据电力客户数据的特点、聚类的目标以及算法的性能等因素综合考虑,选择合适的聚类算法。

### 三、聚类分析在电力客户分类中的应用

#### 1. 基于用电行为的客户分类

电力用户的用电行为反映了用户的特点,对用户用电的时间、电量以及不同时间段内的峰值或谷值等用电行为指标进行聚类分析,能够得到不同类型用户的用电行为分类结果,比如将用户划分为高耗能型、低耗能型、峰时用电型、谷时用电型等。对于高耗能型用户,电力公司应着重关注该类用户是否具有节能降耗的需求,并为其提供相关的咨询服务及技术手段;而对于高峰时用电型用户,则可以通过建立合适的分时电价政策来促进这类用户改变自身的用电习惯并减少高峰期的用电比例,从而达到降低电网运行成本的目的<sup>[4]</sup>。

取某一地区用户作为研究对象,对这一年的电量使用情况统计,并采用K-Means算法对该区域内的客户进行聚类,根据聚类结果划分出三种类型的客户:第一种类型是该区域内用电量最高的客户群体,这类客户的用电情况呈现全天均匀的特点,即每天早晚时段的用电均处于高位状态;第二种类型则是该区域内用电量最低的一批客户,其特点是主要集中在晚间用电;第三种则是在早、晚两个时间段内出现明显用电峰值的客户群体,即日常生活中常见的“两头忙”一族。对于上述三个不同类型的客户群组,供电部门可以根据其各自特点采取针对性的服务措施来提升服务质量与水

平,比如针对第一类客户向其提供相关节能电器产品的推荐方案以及节电指导意见,而对于第二类客户可为其推出夜间用电价格优惠政策或电费折扣活动,为第三类客户提供峰谷电价调整建议等。

#### 2. 基于缴费行为的客户分类

客户的缴费情况也可以反映出用户的信用水平及消费特点。通过用电客户的缴费周期、缴费金额、欠费次数等缴费行为特征的聚类分析结果,可以把用户划分为不同的缴费信用等级以及缴费行为习惯类别。比如可以将用户划分为按时缴费型、拖欠缴费型、预存缴费型等。对于按时缴费型客户,供电企业应给予一定激励政策或优惠措施,以此来增加该类客户与企业的黏合程度;而对于拖欠缴费型客户则需加大催缴力度并建立有效的预警机制,尽量减少欠费发生概率。

对于电力企业的客户交费数据,利用层次聚类方法对该客户的特征属性进行分析后发现。可划分为以下三类:(1)按时缴费型:这部分用户一般都是一次性全额缴齐当月应缴电费;(2)拖欠缴费型:这部分用户的欠款金额较大,并且拖欠的时间比较长;(3)预存缴费型:这部分用户往往能够及时地按照自己的使用量提前预存电费,这种行为说明其有着很强的消费计划意识。基于以上三个不同的客户群体所表现出的不同特性,结合当前我国电力行业的发展现状及发展趋势,建议电力企业实际运营中应该分别采取对应的经营策略与服务措施,如对按时缴费型客户提供优质的服务体验,对拖欠缴费型客户加强信用评估和风险管理,对预存缴费型客户提供预存电费优惠活动等<sup>[5]</sup>。

#### 3. 基于负荷特性的客户分类

电力用户负荷特性的主要参数有:负荷的大小;负荷的变化规律(即日负荷率);负荷曲线上各点的时间分布特征等。根据这些参数,可把用户的负荷分成各种不同的负荷群组,比如工业负荷型、商业负荷型、居民生活负荷型等。对于一个电网来说,不同的负荷类型有不同的作用与意义,在一定程度上反映着该区域的社会经济发展水平及能源利用效率状况,并直接关系到电力系统及其元件的安全性、经济性和稳定性问题。一般而言,工业负荷的特点是其负荷量大而稳定,而且对其所接入的配网线路的要求相对严格一些,以保证高可靠地连续供电。商业负荷是在营业期间内用电量比较大,

但也有相当大的季节性以及节假日的波动情况。居民负荷则表现为在夜间或双休日时用电量较多,呈现比较典型的尖峰负载现象。

以某城市电力客户的负荷作为样本对 DBSCAN 算法进行了聚类分析,得到的结果把客户分成了以下4种:第一类是大型工业客户,负荷功率较大、变动范围小;运行周期较规律;第二类是商业综合体客户,负荷在工作日内的经营时间段内有明显峰值,在休息日及节假日时也有一定幅度的变化;第三类是普通居民用户,负荷有一个比较集中的用电时段(一般是在晚上七点到十点左右);白天负荷较少;第四类是一般的小型工业企业与服务业混合户,负荷特性处于工业和商服两者的中间位置,有一定的波动。根据不同类型的负荷情况来制定合理的发电方案以及网络调峰等措施有利于电力企业的供电管理,能够有效地配置电力资源并使整个电力系统高效稳定地运行。

#### 4. 基于客户价值的客户分类

客户价值(Customer Value)是指由一个或多个客户对企业所做出的经济和社会等方面的贡献减去企业为客户提供的产品和服务所需花费的各种费用后的余额,即客户为企业带来的利润与投入资源之间的差额,是衡量企业和顾客之间关系质量的一个重要参数。通过建立客户用电量、电费缴纳情况以及电价敏感性等特征变量的数据集,利用模糊C均值算法对不同类型的客户群组进行分类分析,从而确定客户的价值并划分出相应的价值级别,包括:高价值客户(High-value Customer, HVC)、中价值客户(Medium-value Customer, MVC)及低价值客户(Low-value Customer, LVC);针对不同类型级别的客户制定差异性的营销策略,例如针对HVC实施个性化服务方案以增强客户满意程度和忠诚度,而针对LVC则可采用降低服务;对于低价值客户,可以采取成本控制措施,优化服务资源配置。

选取某一地区电力公司的客户数据作为样本建立客户价值评价指标体系,主要包括客户的月均用电量(千瓦时)、电费收入占比(元/户·年)、平均交费时间间隔(天)及投诉数(次),通过模糊C均值聚类方法对客户价值分层,将其划分为3个层次:高价值客户(HVH),中价值客户(MVM),低价值客户(LVL);其中,HVH型客户是该区域所有客户中

的前20%,MVM型客户是后75%与HVH型客户之间的部分;LVL型客户是剩余15%;HVH型客户主要表现为月均用电量且稳定,电费缴纳较为及时且无欠缴情况发生,同时用户投诉较少甚至没有,因此其对电网公司有较高的忠诚度并能带来较大的收益;MVM型客户整体上表现一般,各指标基本介于上述两组之间,在某种程度上具有一定的代表性;LVL型客户则表现出用电量小,存在欠费现象,并伴有一定程度上的投诉行为。根据以上分析可知,HVH型客户应享受最高等级的服务,而LVL型客户则需要更多的关注和服务支持。在此基础上可分别针对这三种类型的不同需求制定相应的差异化服务策略,比如对于HVH型客户给予专人负责制、提供优质的个性化服务以及确保对其供电可靠性要求更高一些;而对于LVL型客户则应在保证正常服务的基础上尽量降低成本。

#### 结束语

聚类分析作为电力客户分类的有效工具,已在电力行业展现出显著价值,从理论原理的阐述到多种算法的应用,再到实际问题的解决,都为电力企业优化服务、提升竞争力提供了有力支撑。伴随着我国电力行业的快速发展和数字化进程加快,电力行业产生的数据量会越来越多且越来越复杂,而作为数据分析方法之一的聚类分析技术也将会不断发展和完善,并且在未来的发展中将会同人工智能、大数据等相关领域进行深度融合以提升电力客户的细分程度以及电力企业的智能化服务水平,从而为促进我国电力行业的可持续发展做出更大的贡献

#### [参考文献]

- [1]何鑫,卜乐,宋恩博,等.基于聚类分析的电力客户分群及特征研究[J].企业改革与管理,2018,(17):95-97.
- [2]韦奇均,孙嘉淼,陆溪源,等.基于电力大数据的电力客户行为分析系统[J].广西电力,2024,47(04):37-43.
- [3]龙鹏,徐涛,吴新瑞,等.基于营销大数据的售电市场客户分类识别研究[J].电气技术,2024,25(10):30-35+41.
- [4]蒋迪.农村用电客户的分类及差异化服务[J].农村电工,2019,27(08):13.
- [5]余锦河,王建波,张思雨,等.电力客户细分服务管理建设与成效[J].电气技术与经济,2023,(02):195-198.