2021 年 第 21 卷 第 35 期
2021, 21(35): 15130-06

科 学 技 术 与 工 程
Science Technology and Engineering

ISSN 1671—1815
CN 11—4688/T

引用格式:刘云天, 孙晶琪, 牛东晓, 等. 基于聚类算法的电力制造业 ERP 数据分类[J]. 科学技术与工程, 2021, 21(35): 15130-15135.
Liu Yuntian, Sun Jingqi, Niu Dongxiao, et al. ERP data classification of power manufacturing industry based on clustering algorithm[J]. Science Technology and Engineering, 2021, 21(35): 15130-15135.

基于聚类算法的电力制造业 ERP 数据分类

刘云天[1], 孙晶琪[1*], 牛东晓[1], 张焕粉[2]
(1. 华北电力大学经济与管理学院, 北京 102206; 2. 北京清畅电力技术股份有限公司, 北京 100085)

摘 要 鉴于传统制造业企业仓储分类方法分类边界不明确、难以适应现代企业信息化发展趋势的问题,提出了基于聚类算法的 ABC 库存分类算法,提升了传统库存分类模型的分类精度和效率。以某电力制造业企业的库存产品数据为研究对象,结合聚类算法与 ABC 分类法提出了适用于实际研究场景的评价函数,并给出了新的基于 *K*-means 算法的 ABC 分类法。采用该分类模型对库存环网柜产品进行分类,并基于企业资源计划(enterprise resource planning,ERP)系统中存储的产品数据将库存环网柜产品分为 A、B、C 三类,根据分类结果对不同类别的环网柜产品采取不同的库存控制策略。研究结果表明,将数据挖掘算法应用于库存管理实现了企业库存管理决策的科学化和智能化。

关键词 电力制造业企业; 聚类算法; ABC 分类法; ERP 系统
中图法分类号 TP301.6; **文献标志码** A

ERP Data Classification of Power Manufacturing Industry Based on Clustering Algorithm

LIU Yun-tian[1], SUN Jing-qi[1*], NIU Dong-xiao[1], ZHANG Huan-fen[2]
(1. School of Economics and Management, North China Electric Power University, Beijing 102206, China;
2. Beijing Qingchang power Technology Co., Ltd, Beijing 100085, China)

[**Abstract**] In view of the unclear classification boundary of the traditional manufacturing enterprise warehouse classification method and the difficulty of adapting to the development trend of modern enterprise informatization, the ABC inventory classification algorithm based on the clustering algorithm is proposed to improve the classification accuracy and efficiency of the traditional inventory classification model. Taking the inventory product data of an electric power manufacturing company as the research object, combining the clustering algorithm and the ABC classification method, an evaluation function suitable for actual research scenarios is proposed, and a new ABC classification method based on the k-means algorithm is given. This classification model is used to classify the inventory ring network cabinet products, and based on the product data stored in the enterprise resource planning(ERP) system, the inventory ring network cabinet products are divided into three categories: A, B, and C. According to the classification results, different categories of ring network cabinet products are taken Different inventory control strategies. The research results show that the application of data mining algorithms to inventory management has realized the scientific and intelligent decision-making of enterprise inventory management.
[**Keywords**] electric power manufacturing enterprises; clustering algorithm; ABC classification; ERP system

伴随着大数据和信息化时代的来临,传统的库存控制管理方法渐渐难以适应现代化企业库存管理的需求。制造业企业往往存在着库存产品繁多,难以管理的问题,对于高经济价值的库存产品进行重点管理,而忽略低经济价值库存产品的管理控制,从长远来看会造成巨大的经济损失,制造业企业如何管理这些种类繁多的库存产品,使库存更加合理已成为亟待解决的难题[1]。

目前,许多学者已经开展了对于制造业企业库存分类规划管理的研究,通过将综合评价方法引入传统分类模型,以实现更准确的库存分类。周卫标[2]等以 *X* 家电企业为例,采用传统的 ABC 分类法融入多重指标法再次分类,优化分类方案,使库存管理达到最优化,使货品货物安排更合理、更科学。Yiğit 等[3]采用三阶段多准则库存分类(multi-criteria inventory classification,MCIC)结合层次分析法以

收稿日期: 2021-09-07; **修订日期**: 2021-10-09
基金项目: 国家重点研发计划(2020YFB1707801)
第一作者: 刘云天(1996—),男,土家族,湖北宜昌人,硕士研究生。研究方向:工业工程、技术经济学。E-mail:liuyuntian@ncepu.edu.cn。
* **通信作者**: 孙晶琪(1974—),女,汉族,黑龙江佳木斯人,博士,副教授。研究方向:电力市场、复杂系统理论、能源经济。E-mail:sjq@ncepu.edu.cn。

及模糊 C 均值算法提出了层次分析-模糊 C 均值(analytic hierarchy process, fuzzy C-meas)多准则 ABC 库存分类方法,通过将该方法与其他方法进行比较,验证了该方法的优越性。于赫等[4]以经济订货批量(economic order quantity, EOQ)模型为基础,结合 A 公司库存具体数据,根据模型计算得出 A 公司库存元件的经济采购量和经济采购批次,通过库存成本的降低显著提高了 A 公司的净利润。方忠民等[5]将传统 ABC 分类法与 FSN(依据物料消耗速度进行分类的库存管理方法)分类法相结合,并根据模糊理论建立价值和消耗量模糊因子,构建了模糊 ABC-FSN 分类法对企业库存物料进行分类管理。Hajiagha 等[6]提出了一种混合模糊-随机多准则 ABC 库存分类方法,利用混合算法求解所提出的模型,将物品分类为极重要(A 类)、中等重要(B 类)和相对不重要(C 类),确定了物品的最小库存水平,降低了需求信息的不确定性。

以上方法均采用传统分类方法对企业库存产品进行管理,存在着分类标准较为单一,在实际应用时存在盘点清查困难等问题。考虑到传统分类方法的局限性,很多学者已经开始探索新的数据分类方法。梁淑蓉等[7]基于权重搜索树提出了改进 K 近邻的高维分类算法,该算法对比 K 近邻分类算法、决策树和支持向量机算法,显著减少了数据分类的时间,相较其他算法也具有更高的分类准确率。张扬等[8]研究了一种新的证据分聚类融合算法。显著提高了分类准确度。王刚等[9]提出了一种新的云计算环境下自适应分类方法,通过提取的分布式语义文本特征并进行验证,发现该方法具有更高的分类精度和效率。从以上研究可知,学者们通过将数据挖掘技术应用到数据分类中,大大提升了数据分类的效率和准确性。

由于传统库存分类方法存在不能准确分类,存量控制困难等问题,现提出基于 K-means 聚类算法的 ABC 库存分类方法。首先从供应商的库存管理系统中提取产品销售数据,然后采用聚类分析算法,对原始数据进行数据预处理,在剔除掉无关数据后,根据各样本数据与各聚类中心的距离对样本数据进行分类,建立数据分类模型对库存产品进行分类,以指导企业的库存管理决策。通过将数据挖掘技术与传统库存管理方法相结合,达到建立科学而有效的库存分析系统的目的。

1　ABC 分类法和 K-means 聚类算法的理论基础

1.1　ABC 分类法

ABC 分类法是将事物在价值、数量等方面的特征作为依据,进行划分和归类、区分重要和一般,以采取不同管理策略的一种方法[10]。ABC 分类法的思想实质上可以这样概括:多数的经济价值往往集中于少数产品,而多数产品仅仅占据少数经济价值,企业应该重视对前者的管理[11]。ABC 分类法的一般步骤如下。

(1)数据获取。从 ERP 管理系统中获取产品的种类、型号、销售额等数据。

(2)分析计算。将从 ERP 管理系统中获取的数据进行分析、计算。继而计算各产品的销售量与销售累计量占总销售量的百分比、各产品销售额与累计销售额占总销售额的百分比。

(3)确定分类标准。根据 ABC 分类法基本思想,把累计销售额为 60% ~80% 的产品归为 A 类,把累计销售额为 20% ~30% 的商品归为 B 类,其余归为 C 类。

(4)依据上一步的分类结果,制订 ABC 分类管理标准表,对 A、B、C 三类对象采取不同的管理策略。

1.2　K-means 聚类算法

聚类分析是数据挖掘中十分热门的一个研究方向,与其他学科的研究领域有着很大的交叉性[12]。可以通过使用聚类分析算法发现各种数据的隐含结构,对数据进行自动分类,进而获知数据的大致分布,能够在诸多领域为各种不同的企业提供决策支持[13]。

K-means 算法是一种得到最广泛使用的聚类算法,其主要思想是:首先通过计算得出各聚类子集内所有数据样本的均值,并将其作为该聚类的代表点,然后把每个数据点归入最邻近的类别中,使评价聚类性能的准则函数趋于最优、从而使同类别中的对象相似度(相似度为一个簇中对象的平均值)较高,而不同类之间的对象相似度较小[14]。算法的一般步骤为:在样本中随机抽取 K 个对象,假设每个对象的初始值等于该簇的平均值,对于剩余对象,根据其与各个簇的质心的距离逐个分配到最邻近的簇,然后重新计算各个簇的平均值,不断重复此过程直至准则函数收敛。

K-means 算法的一般步骤如下。

假定存在数据点集 $D = \{x_1, x_2, \cdots, x_n\}$,以及 k 个簇 $C = \{C_1, C_2, \cdots, C_k\}$,每个簇的聚类中心为

$$\overline{X}_i = \frac{1}{n_{C_i}} \sum_{x \in C_i} X \tag{1}$$

式(1)中:n_{C_i} 表示第 i 个簇中数据点 x 的个数。

构建准则函数为

$$E = \sum_{i=1}^{i} \sum_{x \in C_i} |X - \overline{X}_i|^2 \tag{2}$$

(1)为数据点集中的每一个数据点随机分配 k 个簇各自的聚类中心 $\bar{x}_i(1 < i \leqslant k)$。

(2)计算每个数据点与每个聚类中心 $\bar{x}_i$ 的距离 $d_i = |\bar{X}_i - X_j|(1 < i \leqslant k)$,将这些数据点根据最小距离原则分配至最邻近聚类。

(3)依据已经划分的数据点重新计算每个聚类的聚类中心 $\bar{x}_i$。

(4)循环步骤(2)和步骤(3),直到准则函数 E 的数值保持不变。

(5)分类完成,得到 k 个聚类。

2 基于 *K*-means 聚类算法和 ABC 分类法的电力制造业产品细分建模

在对库存产品进行分类时,首先为每一个类选择一个聚类中心,依据 ABC 分类方法将库存产品数据分为三个簇,其中 A 类库存产品是指价值 80% 销售额的占据 20% 数量的产品型号,假设该类产品的总销售额为 M,总库存数量为 N,则定义 A 类产品的平均销售额 C 作为 A 类产品的中心为

$$C = \frac{0.8M}{0.2N} \tag{3}$$

定义 φ 为评价函数为

$$\varphi = \frac{Q(i)}{C} \tag{4}$$

式(4)中:$Q(i)$ 为 i 型号产品的销售额。

在 *K*-means 算法中,选取合适的评价函数非常重要,因此在本文所构建的模型中将评价函数的选取与 ABC 分类法相关联,通过设置合适的评价函数来衡量库存产品 i 的销售额与本类产品销售额的聚类中心间的距离。规定若 $\varphi < 0.8$,则表示该产品的销售额与聚类中心距离过远,并将该产品进行次一级分类评价;反之则将该产品归入当前类别。

将 *K*-means 算法中的评价函数引入分类算法,解决了 ABC 分类法对于分类边界判断模棱两可的问题。假设 A 类产品的销售额占比为 80%,若仓库中存在 m 产品,库存占比为 65%,则 m 产品库存占比与销售额占比之和未达到 80%,根据 ABC 分类法的理论,m 产品属于 A 类。但假定 m 产品的销售额很少,那么将 m 产品归入 A 类是不合理的。因为在模型中引入评价函数,显然会使 φ 的值变小,从而 m 产品不应该归入 A 类。这种改进方法避免了传统方法的边界分类不准确性,使得分类边界更加清晰,加强了对库存产品分类的灵活性和准确性。

基于 *K*-means 算法的 ABC 分类法的步骤如下。

(1)根据各型号环网柜比例和基于 ABC 分类法得出的分类,计算各类库存环网柜的聚类中心 C_i。

(2)对于每种商品 x。

(3)计算评价函数:$\varphi = \frac{Q(i)}{C}$。

(4)如果 $\varphi > R$(R 为当前类别环网柜评分比例),则将 x 归入当前类别。

(5)在当前品类产品销量比例的基础上,增加环网柜 x 占总销量的比例。

(6)否则转换为 ABC 中的下一类别,循环步骤(4)~步骤(6)。

(7)结束判断。

(8)结束循环。

(9)计算新的聚类中心 C_i。

(10)循环步骤(2)~(9),直至聚类中心 C_i 保持不变。

通过在计算机中运行上述算法,可以获得环网柜分类信息表,在表中可以直观的观察到各类别库存环网柜的销售额在总销售额中占多大比例。根据嵌套聚类算法后的环网柜分类结果,库存中心可以针对不同类别的库存环网柜产品采取不同的库存管理策略。比如对于 A 类环网柜,采购需求量较为巨大,必须保证这些型号的环网柜在多个仓库中的储存量,不允许出现缺货的情况;对于 B 类环网柜,可以采取适当控制的策略,应该在可能的范畴内尽量减少库存;对于 C 类中的各型号环网柜产品,只需采用经济批量即可。

3 实证计算

3.1 问题提出

某电力制造业企业库存中心每天要处理大量各种型号的环网柜,对于这些不同型号、具有不同经济价值的环网柜产品,考虑到客户对环网柜产品的多样化需求以及中心仓库较为有限的库存容量,企业库存中心必须根据各型号环网柜的月度销量,对不同型号的环网柜产品采取不同的库存策略(比如计划库存量、经济订货量),从而提高产品出入库的工作效率、刺激环网柜产品的销售并尽可能的减少库存。在算例中需要解决的问题是:对于该企业的环网柜库存中心而言,通过数据挖掘算法挖掘出哪些型号的环网柜是占据库存空间较小,但是最有经济价值的,则应该对这些型号的环网柜产品进行重点管理。因此,可以依据该算法,利用计算机得出各环网柜型号的从属分类,企业的库存中心可以依据分类结果对不同型号的

环网柜采取不同的管理策略,以实现最佳的库存管理效果。

3.2　数据采集

本文中数据来源于某电力制造业企业的后台 ERP 数据库中的历史环网柜产品销售表。根据算法输入的要求,本次建模需要的数据包括环网柜的型号(即库存品名称),环网柜的月度销售量,环网柜的月度销售额。因此从 ERP 系统中导出的环网柜销售记录表如表 1 所示。

表 1　环网柜销售记录表(部分)

Table 1　Ring network cabinet sales record sheet (part of the table)

产品序号	库存品名称	数量/件	销售额/元
1	AC10 kV	6	789 160.32
2	AC10 kV,125 A	2	28 847.52
3	AC10 kV,630 A	1	22 632.48
4	AC10 kV,630 A	3	34 510.32
5	AC10 kV,630 A	2	31 262.40
6	QLG-12	12	937 440.00
7	QLG-12	2	74 880.00
8	QLG-12	6	179 280.00
9	AC10 kV,2,2	2	155 244.96
10	AC10 kV,2,4	3	372 621.60
11	AC10 kV,630 A	6	104 850.72
12	AC10 kV,630 A	10	130 291.20
13	AC10 kV,125 A	6	98 897.76
14	AC10 kV,125 A	2	32 965.92
15	AC10 kV,630 A	2	26 058.24
16	AC10 kV,630 A	2	34 950.24
⋮	⋮	⋮	⋮

3.3　数据处理

从 ERP 管理系统的数据库中导出的原始数据存在着很多空缺和不一致的质量不高的数据,在建模分析时这些数据会对分析结果造成不良影响,因此在建模计算前必须对数据进行加工处理,一般的处理步骤大致包括数据清理、数据集成、数据变换和数据归约[15]。因为在本文的实证研究中以月为时间跨度,需要根据产品的月度销售额数据来对产品数据进行聚类,但从数据库中导出的数据表只有每批次销售额这一属性,因此需要对环网柜销售记录表中的库存数据进行处理。首先将原始数据中的空缺值等低质量数据去掉,然后根据环网柜销售记录表的库存品名称、每批次销售量和每批次销售额数据,通过运行代码计算得到月销售额数据,然后根据分类算法通过计算机计算得到相应的分类。

3.4　数据计算及分类

利用传统的 ABC 分类法对原始库存数据表进行数据分类的结果如表 2 所示,划分的库存环网柜累计金额占比趋势变化如图 1 所示。

根据累计金额占比趋势图可知,17 类不同型号的环网柜中,有 3 类产品被分为 A 类,4 类产品被分为 B 类,10 类产品被分为 C 类。然后将传统方法分类后的数据代入改进的基于聚类算法的 ABC 分类模型。在利用 *K*-means 算法进行聚类时,聚类中心的变化如表 3 所示。

在 A、B、C 三类的聚类中心不再改变后,利用 *K*-means 聚类算法进行二次分类的结果如表 4 所示,累计金额占比变化趋势如图 2 所示。

表 2　数据分类表

Table 2　Data classification table

序号	库存品型号	累计销售额/元	累计百分比	类别
1	AC10 kV,2 630 A	7 902 637.44	0.553	A
2	QLG-12/630	10 792 373.64	0.755	A
3	WKXGN15-12CCVVVV	11 584 373.64	0.811	A
4	QLG-12/630-5L	12 009 173.64	0.840	B
5	QLG-12/630-2L2F	12 292 373.64	0.860	B
6	QLG-12/630-6D1	12 535 445.64	0.877	B
7	QLG-12/630-4L2D6	12 772 245.64	0.894	B
8	QLG-12/630-6D1	12 991 413.64	0.909	C
9	QLG-12/630-3L1F4-1-3	13 170 693.64	0.922	C
10	QLG-12/630-4L6D1	13 343 493.64	0.934	C
11	QLG-12/630-1L2D + KPT	13 503 802.54	0.945	C
12	QLG-12/630-1L4D1Lpt * 2	13 663 012.54	0.956	C
13	QLG-12/630-4D1L	13 821 862.54	0.967	C
14	QEFGX-12/630-LD	13 953 262.54	0.976	C
15	QLG-12/630-3L3D6-1-5	14 080 462.54	0.985	C
16	AC10 kV,630 A	14 199 064.54	0.994	C
17	QLG-12/630-1L1F2	14 289 334.54	1.000	C

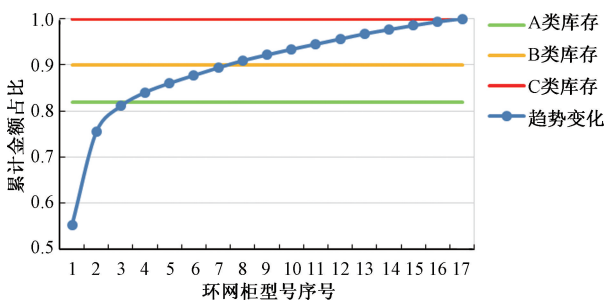

图 1　ABC 分类法下的累计占比库存金额变化趋势图

Fig. 1　Change trend chart of cumulative proportion of inventory amount under ABC classification

表 3　聚类中心变化表

Table 3　Cluster center change table

初始聚类中心	最终聚类中心
C_1:44 727.31	C'_1:42 997.50
C_2:91 374.77	C'_2:135 200.00
C_3:48 938.35	C'_3:53 027.00

表 4 二次数据分类表

Table 4 Secondary data classification table

序号	库存品型号	累计销售额/元	累计百分比	类别
1	AC10 kV,2 630 A	7 902 637.44	0.553	A
2	QLG-12/630	10 792 373.64	0.755	A
3	WKXGN15-12CCVVVV	11 584 373.64	0.811	B
4	QLG-12/630-5L	12 009 173.64	0.840	B
5	QLG-12/630-2L2F	12 292 373.64	0.860	C
6	QLG-12/630-6D1	12 535 445.64	0.877	C
7	QLG-12/630-4L2D6	12 772 245.64	0.894	C
8	QLG-12/630-6D1	12 991 413.64	0.909	C
9	QLG-12/630-3L1F4-1-3	13 170 693.64	0.922	C
10	QLG-12/630-4L6D1	13 343 493.64	0.934	C
11	QLG-12/630-1L2D + KPT	13 503 802.54	0.945	C
12	QLG-12/630-1L4D1Lpt * 2	13 663 012.54	0.956	C
13	QLG-12/630-4D1L	13 821 862.54	0.967	C
14	QEFGX-12/630-LD	13 953 262.54	0.976	C
15	QLG-12/630-3L3D6-1-5	14 080 462.54	0.985	C
16	AC10 kV,630 A	14 199 064.54	0.994	C
17	QLG-12/630-1L1F2	14 289 334.54	1.000	C

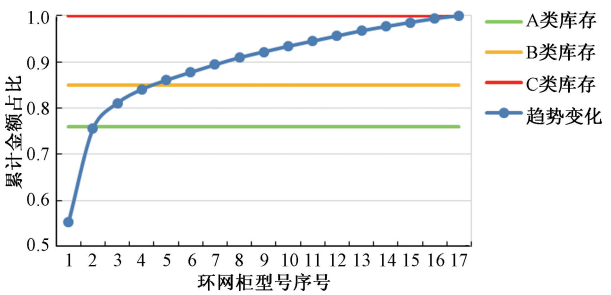

图 2 基于聚类分类算法的累计占比库存金额变化趋势图

Fig. 2 The trend chart of the cumulative inventory amount based on the cluster classification algorithm

结果表明,AC10 kV、2 630 A 和 QLG-12/630 这两类型号的环网柜属于 A 类产品,销售额占比约 76%,WKXGN15-12CCVVVV 和 QLG-12/630-5L 型号的环网柜属于 B 类产品,销售额占比 8%,其余产品属于 C 类,销售额占比 16%。故应该对 AC10 kV、2 630 A 和 QLG-12/630 型号的环网柜进行重点控制,应该严格控制其计划与采购、库存储备量、订货量和订货时间;对 WKXGN15-12CCVVVV 和 QLG-12/630-5L 型号的环网柜进行适当控制,在需求范围内适当减少库存;对于其余的 C 类产品可以放宽控制,采用经济批量。

4 结论

采用 *K*-means 聚类算法和 ABC 分类法建立了库存分类模型,并对某电力制造业企业生产的多种型号环网柜进行了产品库存分类,得到以下结论。

(1)在 ABC 分类法中引入评价函数较好的解决了传统分类的弊端,这种改进方法避免了传统方法在边界分类中的不准确性,通过定义合适的评判标准使得评价函数可以对各型号的库存环网柜产品与各聚类中心的距离进行定量的判断,也加强了各类环网柜划分的灵活性,增强了算法的准确性。

(2)根据数据挖掘聚类算法与经典库存分类方法建模计算得出的分类结果来指导企业库存决策,对于企业库存管理有着较强的指导意义。数据挖掘技术与 ERP 信息系统的紧密结合大大增强了 ERP 的分析和决策功能,充分满足了用户的需求,利用挖掘得到的 ERP 数据除了可以进行产品库存分类、股利分配决策、企业价值评估和财务预警外,还能为许多其他的企业决策提供支持。

参 考 文 献

[1] 刘佳兴. 基于聚类分析的库存分析系统研究与开发[D]. 沈阳: 沈阳工业大学, 2017.
Liu Jiaxing. Research and development of inventory analysis system based on cluster analysis[D]. Shenyang: Shenyang University of Technology, 2017.

[2] 周卫标, 梁碧珊, 高 敏. 基于 ABC 分类法指标扩展的库存分类管理分析[J]. 现代商贸工业, 2018, 39(7): 64-66.
Zhou Weibiao, Liang Bishan, Gao Min. Analysis of inventory classification management based on ABC classification index expansion [J]. Modern Business and Trade Industry, 2018, 39(7): 64-66.

[3] Yiğit F, Esnaf Ş. A new fuzzy *C*-means and AHP-based three-phased approach for multiple criteria ABC inventory classification [J]. Journal of Intelligent Manufacturing, 2021, 32(6): 1517-1528.

[4] 于 赫, 王忠吉, 左圆圆. 基于 EOQ 模型的 A 公司库存管理优化研究[J]. 物流工程与管理, 2016, 38(7): 106-107, 109.
Yu He, Wang Zhongji, Zuo Yuanyuan. Research on a company's inventory management optimization based on EOQ model[J]. Logistics Engineering and Management, 2016, 38(7): 106-107, 109.

[5] 方忠民, 韩福义, 马 蓉. 模糊 ABC-FSN 分类法在企业库存管理中的应用[J]. 物流技术, 2019, 38(1): 114-119.
Fang Zhongmin, Han Fuyi, Ma Rong. The application of fuzzy ABC-FSN classification method in enterprise inventory management [J]. Logistics Technology, 2019, 38(1): 114-119.

[6] Hajiagha S H R, Daneshvar M, Antucheviciene J. A hybrid fuzzy-stochastic multi-criteria ABC inventory classification using possibilistic chance-constrained programming[J]. Soft Computing, 2021, 25(2): 1065-1083.

[7] 梁淑蓉, 陈基漓, 谢晓兰. 基于权重搜索树改进 *K* 近邻的高维分类算法[J]. 科学技术与工程, 2021, 21(7): 2760-2766.
Liang Shurong, Chen Jili, Xie Xiaolan. Improved *K*-nearest neighbor high-dimensional classification algorithm based on weighted search tree[J]. Science Technology and Engineering, 2021, 21(7): 2760-2766.

[8] 张 扬, 杨建华, 侯 宏. 一种基于证据理论的数据分聚类融合算法[J]. 科学技术与工程, 2018, 18(1): 54-58.
Zhang Yang, Yang Jianhua, Hou Hong. A data clustering fusion algorithm based on evidence theory[J]. Science Technology and En-

gineering, 2018, 18(1): 54-58.

[9] 王 刚, 杨 波, 杨明杰. 云计算环境下分布式语义文本自适应分类方法[J]. 科学技术与工程, 2018, 18(7): 208-212.
Wang Gang, Yang Bo, Yang Mingjie. Distributed semantic text adaptive classification method in cloud computing environment[J]. Science Technology and Engineering, 2018, 18(7): 208-212.

[10] 刘培永, 陆志强. 多维度 ABC 分类法在原材料库存控制中的应用[J]. 机械制造, 2017, 55(2): 99-102.
Liu Peiyong, Lu Zhiqiang. The application of multi-dimensional ABC classification in raw material inventory control[J]. Machinery Manufacturing, 2017, 55(2): 99-102.

[11] 胡靖枫, 何利力, 周庆燕. 基于聚类分析的 ABC 库存分类方法研究[J]. 工业控制计算机, 2015, 28(3): 147-148.
Hu Jingfeng, He Lili, Zhou Qingyan. Research on ABC inventory classification method based on cluster analysis[J]. Industrial Control Computer, 2015, 28(3): 147-148.

[12] 陶 莹, 杨 锋, 刘 洋, 等. *K* 均值聚类算法的研究与优化[J]. 计算机技术与发展, 2018, 28(6): 90-92.
Tao Ying, Yang Feng, Liu Yang, et al. Research and optimization of *K*-means clustering algorithm[J]. Computer Technology and Development, 2018, 28(6): 90-92.

[13] 盘俊良, 石跃祥, 李娉婷. 一种新的粒子群优化聚类方法[J]. 计算机工程与应用, 2012, 48(8): 179-181.
Pan Junliang, Shi Yuexiang, Li Pingting. A new particle swarm optimization clustering method[J]. Computer Engineering and Applications, 2012, 48(8): 179-181.

[14] 程俊文, 李慧娟, 曹志强. 基于 *K*-means 算法和用电信息采集系统的防窃电研究[J]. 供用电, 2019, 36(1): 75-80.
Cheng Junwen, Li Huijuan, Cao Zhiqiang. Research on anti-stealing electricity based on *K*-means algorithm and electricity information collection system[J]. Power Supply, 2019, 36(1): 75-80.

[15] 张治斌, 刘 威. 浅析数据挖掘中的数据预处理技术[J]. 数字技术与应用, 2017(10): 216-217.
Zhang Zhibin, Liu Wei. Analysis of data preprocessing technology in data mining[J]. Digital Technology and Application, 2017(10): 216-217.