

基于主成分分析法的制造业信息化绩效聚类分析 ——以沈阳为例

杜晓君¹, 杨 雷^{1,2}

(1. 东北大学 工商管理学院, 辽宁 沈阳 110819; 2. 沈阳师范大学 旅游管理学院, 辽宁 沈阳 110034)

摘 要: 制造业信息化绩效是评价制造业综合运营能力、客户响应速度以及创新能力的重要指标。结合理论研究和企业调研, 提出了制造企业信息化水平评价指标体系, 并采用主成分分析法建立了沈阳市工业企业信息化水平的数学评价模型。最后, 采用 K -means 算法对 100 家沈阳市制造企业进行了信息化水平聚类分析, 认为从强调规模向强调效率、综合运营能力和综合盈利能力转变, 注重软件和人才, 以契合企业战略目标作为信息化实施的根本准则, 是提升制造企业的核心竞争力的根本途径。

关 键 词: 制造业; 信息化; 工业化; 主成分分析; K -means 算法

中图分类号: G 203

文献标志码: A

文章编号: 1005-3026(2013)07-1053-04

Cluster Analysis of Manufacturing Informatization Performance Based on the Principal Component Analysis: Taking Shenyang as an Example

DU Xiao-jun¹, YANG Lei^{1,2}

(1. School of Business Administration, Northeastern University, Shenyang 110819, China. 2. School of Tourism & Hospitality Management, Shenyang Normal University, Shenyang 110034, China. Corresponding author: YANG Lei, E-mail: yanglei2000@126.com)

Abstract: Manufacturing informatization performance is an important index for evaluating the comprehensive operation ability, customer response speed and innovation ability of manufacturing industry. Combining theoretical study and enterprise investigation, an evaluation index system of informatization level was established and a mathematical evaluation model for Shenyang's industrial enterprises' informatization level, based on the principal component analysis, was set up. Lastly, using the K -means algorithm, a cluster analysis of 100 Shenyang's industrial enterprises was carried out, and it was concluded that for manufacturing, the basic approaches to enhance its core competitiveness are shifting from size to efficiency, comprehensive operation ability and comprehensive profitability, giving priority to software and talents, and taking enterprises' strategic target as the fundamental principle for informatization.

Key words: manufacturing; informatization; industrialization; principal component analysis; K -means algorithm

制造业信息化既是实现国民经济和社会信息化的基础,也是用高新技术提高工业整体素质和竞争力的重要手段,同时也是促进工业实现快速、持续发展的重要支撑^[1-4]。制造业信息化的途径通常包括推进信息技术在产品研发设计、生产过

程控制、经营管理、营销流通等各个环节的应用、渗透和融合,推动装备制造业与生产性服务业的融合互动发展以及战略性新兴产业的培育。

本文依据国家确立的信息化体系框架,考虑我国制造业的发展阶段以及在信息化方面的特

收稿日期: 2013-01-06

基金项目: 国家社会科学基金资助项目(04BJY004); 教育部人文社会科学基金资助项目(07JA790081); 辽宁省发改委基金资助项目(LF2012YTH001)。

作者简介: 杜晓君(1964-),女,辽宁盖州人,东北大学教授,博士生导师。

点,建立了能够较全面体现制造业信息化水平的评价指标体系.为消除多个评价指标间的重叠信息,简化工业企业信息化水平的聚类分析,应用主成分分析法从具有一定相关性的多个评价指标中提取出能反映沈阳市制造企业信息化水平的 8 个综合评价指标,并构建了制造企业信息化水平评价数学模型.

1 测算方法选择与指标体系构建

1.1 测算方法选择

现有的指标综合评价方法主要有波拉特法、层次分析法、模糊综合评价法、数据包络法、专家评分法和灰色关联分析等方法^[5-7].沈阳市工业企业信息化水平评价受到诸多可观指标因子的影响,多个可观指标间存在一定相关性,彼此之间不是相互独立的.主成分分析法是通过降维技术把多个变量转化为少数几个主成分的统计分析方法,在具有一定相关性的多指标评价方面取得了良好的效果^[8-9].

1.2 指标体系构建

构建了能系统地描述工业企业信息化水平的指标体系,该指标体系由 2 个一级指标和 26 个二级指标构成,如表 1 所示.

2 综合评价指标提取

2.1 标准化处理

由于各评价指标之间的量纲不同,在利用主成分分析法进行分析前需要对原始数据进行标准化

处理.本文采用 Z 分数法,利用公式 $x'_{ij} = \frac{x_{ij} - x_j}{s_j}$ 对

原始数据进行标准化处理,其中 x'_{ij} 是标准化处理后数据, x_j 和 s_j 分别为第 j 个指标的均值和方差.

2.2 综合评价指标提取

采用 SPSS 计算得到相关系数矩阵的特征值、贡献率和累计贡献率的关系,如表 2 所示.

由表 2 得出,SPSS 软件的前 8 个综合评价指标累计贡献率为 77.024%,可以用来构建信息化水平评价数学模型.

2.3 综合评价指标命名

通过分析旋转后的因子载荷矩阵,得到综合评价指标和原指标之间的关系.例如,综合评价指标 1 在已接入内部办公网的计算机总数、企业拥有计算机总数、工业基础设施设备台数、信息化项目投资总额、信息化咨询及实施费用总额、信息系

表 1 工业企业信息化水平评价指标

Table 1 The evaluation index system of industrial enterprises' informatization level

一级指标	二级指标
信息化基础	信息化投入比例 X_1
	企业人数 X_2
	与信息化工作相关的人员总数占企业员工总数比例 X_3
	已接入内部办公网的计算机总数 X_4
	企业拥有计算机总数 X_5
	工业基础设施设备台数 X_6
	数控设备台数 X_7
	企业销售收入总额 X_8
	信息化项目投资总额 X_9
	信息化咨询及实施费用总额 X_{10}
	信息系统软件投资总额 X_{11}
	信息系统硬件投资总额 X_{12}
	项目规划资金投入额 X_{13}
信息化贡献	主导产品研发周期 X_{14}
	交货周期提升率 X_{15}
	主导产品交货周期 X_{16}
	财务决算速度提升率 X_{17}
	完成财务决算速度 X_{18}
	产品合格率 X_{19}
	客户订单响应速度 X_{20}
	客户订单平均响应速度 X_{21}
	按期交货率 X_{22}
	企业利润总额 X_{23}
	库存占用总额 X_{24}
	企业投入 X_{25}
	企业运转状况 X_{26}

表 2 相关系数矩阵的特征值及方差贡献率和累计方差贡献率

Table 2 Eigenvalue, contribution percentage and cumulative percentage of the correlation matrix

综合评价指标	特征值	贡献率/%	累计贡献率/%
1	6.661	25.619	25.619
2	3.207	12.334	37.952
3	2.628	10.108	48.060
4	2.269	8.727	56.787
5	1.778	6.839	63.626
6	1.321	5.081	68.707
7	1.130	4.348	73.055
8	1.032	3.969	77.024

统软件投资总额这 6 个指标上具有较大载荷.综合以上分析结果,6 个指标主要反映了信息化的

规模,故可以定义综合评价指标 1 为信息化规模指标 f_1 . 仿照综合评价指标 1 的命名方法,分别对其余 7 个综合评价指标进行命名,依次为效率指标 f_2 , 总体收益指标 f_3 , 硬件投入指标 f_4 , 综合运转指标 f_5 , 人员规模指标 f_6 , 产品合格率指标 f_7 以及研发周期指标 f_8 .

2.4 信息化水平评价模型

各综合评价指标的贡献率代表了各综合评价指标在信息化水平评价中所占有的权重,贡献率越大表明该综合评价指标对信息化水平的影响越大. 因此,可基于各主成分与相应贡献率之积的和来求取最终评价价值:

$$F = 0.256f_1 + 0.123f_2 + 0.101f_3 + 0.087f_4 + 0.068f_5 + 0.051f_6 + 0.043f_7 + 0.040f_8.$$

3 制造企业信息化水平聚类分析

3.1 K-means 算法

采用 K-means 算法,初始分类一般随机选择 K 个聚类中心,然后根据样本到聚类中心距离最小的原则对样本进行分类.

假设训练样本是 $\{x^{(1)}, \dots, x^{(m)}\}$, $x^{(i)} \in \mathbf{R}^n$, 算法步骤如下:

步骤 1 初始聚类. 把 m 个样本分成 $w_1, \dots, w_k$ 类, 计算每个聚类的平均值 $u_1, \dots, u_k$ 和每类畸变函数 J_i 和总畸变函数 J .

$$u_i = \frac{1}{N_i} \sum_{y \in w_i} y,$$
$$J_i = \sum_{y \in w_i} (y - u_i)^2,$$
$$J = \sum_{i=1}^k J_i.$$

其中 N_i 为类 w_i 中的样本数目.

步骤 2 修改聚类中心,使总的畸变函数 J 最小. 修改聚类中心的方法是把样本从一个类移动到任一其他类中,如果畸变函数减小则允许移动,否则不允许. 选择样本 $y \in w_i$, 把样本 y 从 w_i 中取出放入 w_j 中, 计算 w_i 和 w_j 畸变函数变化量 ΔJ_i 和 ΔJ_j 及总畸变函数变化量 ΔJ :

$$\Delta J_i = J'_i - J_i,$$
$$\Delta J_j = J'_j - J_j,$$
$$\Delta J = \Delta J_i + \Delta J_j.$$

如果 ΔJ 小于零则允许移动,否则不允许移动. 如果允许移动,修改 u_i 和 u_j 及总的畸变函数 J .

步骤 3 重复步骤 2,迭代至 N 次, J 基本不变,则迭代停止,完成了样本聚类,此时的总畸变函数 J 称为稳态畸变函数 J_s , 否则继续步骤 2.

3.2 制造企业信息化水平聚类分析

采用 3.1 节中的 K-means 聚类算法,取分类数 $K=10$, 将 100 家企业按照综合评价指标划分成 10 类,归类后每类的综合评价指标值由该类中所有工业企业综合评价指标值的平均值来表示,10 类企业的综合评价指标值和最终评价价值如表 3 所示.

表 3 10 类企业综合评价指标及最终评价价值
Table 3 The comprehensive evaluation indexes and final evaluation values of 10-category enterprises

企业类型	f_1	f_2	f_3	f_4	f_5	f_6	f_7	f_8	F
C1	-0.125	0.080	-0.199	-0.251	-0.386	-0.188	0.397	0.008	-0.082 5
C2	0.783	-0.282	1.723	0.984	1.944	0.196	-0.154	0.087	0.564 4
C3	-0.205	-0.050	-0.252	0.005	-0.347	-0.042	-1.844	-0.188	-0.196 2
C4	0.068	-0.104	-0.184	-0.292	-0.141	-0.411	0.548	0.512	-0.025 9
C5	0.375	-0.194	-0.755	0.446	0.524	0.201	0.962	0.024	0.122 9
C6	0.120	-0.276	-0.082	-0.364	1.644	0.249	-0.160	0.023	0.075 4
C7	-0.122	-0.196	-0.024 4	-0.361	3.642	-0.665	-0.001	0.079	0.127 6
C8	-0.505	-0.400	-0.002	2.749	-0.098	-1.443	0.552	0.025	0.005 0
C9	-0.375	-0.454	-0.761	1.002	2.396	-0.408	0.633	0.047	0.029 7
C10	0.223	-0.368	-0.182	-0.521	-0.542	-0.756	1.183	0.091	-0.072 8
均值	0.023 7	-0.224 4	-0.071 8	0.339 7	0.863 6	-0.327	0.212	0.071	0.054 8

从表 3 中分析得出: C1 类企业其特征是在信息化规模、效率、综合运营等方面均低于平均水平; C2 类企业是信息化绩效最优的企业,在信息化规模、综合投入、人员规模、总体收益以及综合

运营等方面均远高于平均水平; C3 类企业产品合格率远低于平均水平,这类企业的信息化应用绩效最差; C4 类企业信息化的应用大大提高了其财务决算速度、客户订单响应速度和按期交货率;

C5 类企业的信息化规模是 10 类企业中最底的, 总体收益最低, 主体产品的研发周期也最长; C6 类企业尽管其信息化硬件投入比例最低, 但是其信息化人员规模最大, 综合运转绩效和收益均略高于平均水平; C7 类企业信息化应用的显著成效是这类企业实现了敏捷反应, 其主导产品交货周期与订单响应速度指标是所有企业中最底的; C8 类企业的信息化绩效不佳: 只在硬件上进行了大量投入, 总体投入不高, 信息化人员规模很小, 综合运作能力最低, 这类企业的存在, 更进一步证实了企业信息化软件和人才的重要性; C9 类企业信息化规模小, 总体收益和企业反应速度均低于平均水平; C10 类企业信息化建设的显著特点是企业效率、产品研发周期和产品合格率是企业中最高的。

4 结 论

1) 数量上占全部调研企业 73% 的企业绩效水平都很低, 表明沈阳的制造企业在以信息化实现转型升级的道路上尚处于从单向应用向集成应用转变的阶段, 整体的信息化绩效不高。

2) 在信息化和工业化融合的实践中, 应尽快从强调规模向强调效率、综合运营能力以及综合盈利能力方向转变。

3) 在企业信息化建设中, 软件和人才至关重要。

要以契合企业战略目标作为信息化实施的根本准则, 才能切实提升企业的核心竞争力。

参考文献:

- [1] Michael E P. Strategy and the Internet [J]. *Harvard Business Review* 2001(3): 63-67.
- [2] Nicholas G C. IT does not matter [J]. *Harvard Business Review* 2003(5): 5-12.
- [3] Donald J B, David J C. Supply chain logistics management [M]. New York: McGraw-Hill Companies 2013: 96-116.
- [4] Anandhi S B. A resource-based perspective on information technology capability and firm performance: an empirical [J]. *MIS Quarterly* 2000 24(3): 169-196.
- [5] 郭金玉, 张忠彬, 孙庆云. 层次分析法的研究与应用 [J]. *中国安全科学* 2008(5): 148-153.
(Guo Jin-Yu, Zhang Zhong-bin, Sun Qing-yun. Study and applications of analytic hierarchy process [J]. *China Safety Science Journal* 2008(5): 148-153.)
- [6] 郑伟丽, 李珊. 基于数据包络法 (DEA) 对甘肃省区域经济增长的分析 [J]. *甘肃科学学报* 2011 23(1): 135-138.
(Zheng Wei-li, Li Shan. Analysis and research of regional economical growth of Gansu: based on DEA [J]. *Journal of Gansu Sciences* 2011 23(1): 135-138.)
- [7] 刘国宏. 中国区域金融中心综合评价研究 [J]. *开放导报*, 2011(3): 40-44.
(Liu Guo-hong. An overview of appraisal of regional financial centers in China [J]. *China Opening Journal* 2011(3): 40-44.)
- [8] Maiyastri K H. Principal component analysis in modelling stock market returns [J]. *Matematika* 2004 20(1): 31-41.
- [9] Jolliffe I T. Principal component analysis [M]. New York: Princeton University Press, 1986: 56-60.
- [9] Ackerman P K, Harris G H, Klimpel R R, et al. Use of xanthogen formats as collectors in the flotation of copper sulfides and pyrite [J]. *International Journal of Mineral Processing* 2000 58: 1-13.
- [10] Li W Z, Qin W Q, Sun W, et al. Electrodeposition of dixanthogen (TETD) on pyrite surface [J]. *Transactions of Nonferrous Metals Society of China* 2007 17(1): 154-158.
- [11] Hope G A, Woods S E, Boyd S E, et al. A SERS spectroelectrochemical investigation of the interaction of butyl ethoxycarbonyl thiourea with copper surface [J]. *Colloids and Surfaces* 2004 232: 129-137.
- [12] Güler T, Hi-Yilmaz C, Gkag G, et al. Adsorption of dithiophosphate and dithiophosphinate on chalcopyrite [J]. *Mineral Engineering* 2006 19: 62-71.
- [13] 李西山, 朱一民. 利用同分异构化学原理研究浮选药剂 Y-89 的同分异构体甲基异戊基黄药 [J]. *湖南有色金属*, 2010 26(2): 19-21.
(Li Xi-shan, Zhu Yi-min. Use the isomeric chemistry principle to study the isomer of flotation reagent Y-89-methyl iso-penty xanthenes [J]. *Hunan Nonferrous Metals* 2010 26(2): 19-21.)
- [3] 刘广义. 硫化铜矿石的综合利用及新型捕收剂研究 [D]. 长沙: 中南大学出版社 2004.
(Liu Guang-yi. Comprehensive utilization of copper sulphide ore and research of new kind of collectors [D]. Changsha: Central South University Press 2004.)
- [4] Targgart A F, Taylor T C, Ince C R. Flotation practice [M]. New York: AIME, 1928: 285-297.
- [5] Marabini A M, Barbaro M, Alesse V. New reagents in sulphide mineral flotation [J]. *International Journal of Mineral Processing* 1991 33: 291-306.
- [6] Israelachvili J N, Adams G E. Direct measurement of long range forces between two mica surfaces in aqueous KNO₃ solution [J]. *Nature* 1976 262: 774-776.
- [7] 朱建光, 朱一民. 浮选药剂的同分异构原理和混合用药 [M]. 长沙: 中南大学出版社 2011.
(Zhu Jian-guang, Zhu Yi-min. The isomeric principle of flotation reagent and mixed reagent [M]. Changsha: Central South University Press 2011.)
- [8] Fu Y L, Wang S S. Neutral hydrocarboxycarbonyl thionocarbamate sulfide collectors: US 4657688 [P]. 1987-04-14.

(上接第 1038 页)