Expert Systems With Applications 50 (2016) 9–16

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Importance–Performance Analysis by Fuzzy *C*-Means Algorithm

Olimpia I. Ban[a], Adrian I. Ban[b,*], Delia A. Tuşe[b]

[a] Department of Economics, University of Oradea, Universităţii 1, Oradea 410087, Romania
[b] Department of Mathematics and Informatics, University of Oradea, Universităţii 1, Oradea 410087, Romania

ARTICLE INFO

Keywords:
Importance–Performance Analysis
Fuzzy clustering
Fuzzy partition
Prototype
Fuzzy *C*-Means Algorithm

ABSTRACT

Traditional Importance–Performance Analysis assumes the distribution of a given set of attributes in four sets, *"Keep up the good work"*, *"Concentrate here"*, *"Low priority"* and *"Possible overkill"*, corresponding to the four possibilities, high–high, low–high, low–low and high–low, of the pair performance–importance. This can lead to ambiguities, contradictions or non-intuitive results, especially because the most real-world classes are fuzzy rather than crisp. The fuzzy clustering is an important tool to identify the structure in data, therefore we apply the Fuzzy *C*-Means Algorithm to obtain a fuzzy partition of a set of attributes. A membership degree of every attribute to each of the sets mentioned above is determined, against to the forcing categorization in traditional Importance–Performance Analysis. The main benefit is related with the deriving of the managerial decisions which become more refined due to the fuzzy approach. In addition, the development priorities and the directions in which the effort of an economic or non-economic entity would be useless or even dangerous are identified on a rigorous basis and taking into account only the internal structure of the input data.

1. Introduction

Importance–Performance Analysis (IPA, for short) is a simple and effective marketing technique which can help practitioners in identifying improvement priorities and direct quality-based marketing strategies. It was proposed by Martilla and James (1977) to analyze the performance of automobile industry and afterwards used in tourism and hospitality industry (Go & Zhang, 1997; Hollenhorst, Olson, & Fortney, 1986), evaluation of restaurants dining (Hu, Chen, & Ou, 2009b), evaluation of tourism policy (Evans & Chon, 1989), identification of the tourists' perceptions of the hotel industry (Lewis & Chambers, 1989), monitoring and improving customer satisfaction (Almanza, Jaffe, & Lin, 1994; Lewis & Chambers, 1989; Martin, 1995), restaurant positioning (Hsu, Byun, & Yang, 1997; Keyt, Yavas, & Riecken, 1994), evaluation of service attributes importance and customer satisfaction (Matzler, Sauerwein, & Heischmidt, 2003), health care (Hawes & Rao, 1985), education (Ortinau, Bush, Bush, & Twible, 1989), slot player experience (Suh & Erdem, 2004), perceptions of dental practices (Nitse & Bush, 1993), mobile telecommunication industry (Pezeshki, Mousavi, & Grant, 2009), etc.

Based on the answers of the customers to the questions about each attribute "How important is it?" and "How well did the product (service) perform?" or based on the answers to the second question and the indirect calculation of importance (see Ban, Ban, & Tuşe, 2015; Feng, Mangan, Wong, Xu, & Lalwani, 2014; Hancock & Klockars, 1991), the performance and importance scores for each attribute are calculated. The points with performance as the first coordinate and importance as the second coordinate are placed on a two-dimensional plot called an IPA grid. The horizontal and vertical axes determine four quadrants and, implicitly, a classification of attributes, as an important step in making strategic marketing decisions, corresponding to the four possible combinations low-high, as in Table 1.

The placement of axes is subject of a continuous debate. They are determined by aggregating the data by arithmetic mean of the values of importance and the values of performances, or they are considered as the mid-points of the scales, or, sometimes, they are placed in a subjective way (see Crompton & Duray, 1985; Kennedy, 1986; Martilla & James, 1977; Mount, 2000; Ortinau et al., 1989). Beside the hidden assumptions in IPA models (see Hu, Lee, Yen, & Tsai, 2009a), the choosing of axes without a natural relationship with the data leads to contradictions, non-intuitive results, subjective interpretations, ambiguities etc., such that many researchers attempted to solve the shortcomings by offering revised versions (see, e.g, Albrecht & Bradford, 1990; Arbore & Busacca, 2011; Bacon, 2012; Caber, Albayrak, & Matzler, 2012; Hu et al., 2009a; Matzler et al., 2003; Oh, 2001). On the other hand, most

* Corresponding author. Tel.:+40 259408656
E-mail addresses: oban@uoradea.ro (O.I. Ban), aiban@uoradea.ro (A.I. Ban), dmerca@uoradea.ro (D.A. Tuşe).

http://dx.doi.org/10.1016/j.eswa.2015.12.023
0957-4174/

Table 1
Quadrants in traditional IPA.

Quadrant	Performance	Importance
1: *Keep up the good work* (GW)	High	High
2: *Concentrate here* (CH)	Low	High
3: *Low priority* (LP)	Low	Low
4: *Possible overkill* (PO)	High	Low

real-world classes are fuzzy rather than crisp. A good example is given even by the attributes in an traditional IPA. We can assume that an attribute characterized by a two-dimensional point (*performance, importance*), especially if it is situated close to axes, has a certain membership degree to two or more quadrants (see Table 1). Bacon (2003) suggests that "a partition of attributes that represents a smoother transition from high to low priorities is more appropriate to improve the validity of establishing priorities in IP space".

Taking into account the above discussion, in the present paper we propose a method of determination of the fuzzy sets, called as in the crisp case, *"Keep up the good work" (GW), "Concentrate here" (CH), "Low priority" (LP), "Possible overkill" (PO),* based on a fuzzy clustering algorithm and corresponding to the quadrants in Table 1. We obtain a fuzzy partition of the initial set of attributes, that is four disjoint fuzzy sets whose union is the initial crisp set of attributes. In fact, a membership degree of each attribute to every set in *{GW, CH, LP, PO}* is determined. This description is more powerful and suitable for deriving managerial decisions than the forcing categorization in the traditional IPA. The efficiency and feasibility of the new method is illustrated on existing data sets in the recent literature. A comparison with the traditional IPA is possible after the defuzzification of the fuzzy partition. Even if we cannot give a mathematical proof of the superiority of the proposed method, the tools which identify natural structure of data – like fuzzy clustering – seem to be more suitable than other artificial approaches in IPA.

2. Fuzzy partition and fuzzy clustering by Fuzzy *C*-Means Algorithm

A fuzzy set on X is a mapping A: $X \to [0, 1]$, where $A(x)$ represents the membership degree of the object x to the fuzzy set A. As usual, we denote the empty fuzzy set by $\varnothing$, that is $\varnothing(x) = 0$, for any $x \in X$ and by X the fuzzy set having the membership degree equal to 1, that is $X(x) = 1$, for any $x \in X$. Two fuzzy sets A and B on X are equal $(A = B)$ if and only if $A(x) = B(x)$, for every $x \in X$. The union $(A \cup B)$ and the intersection $(A \cap B)$ of two fuzzy sets A and B may be defined in many ways, based on triangular conorms and triangular norms, respectively. In this paper we consider the following definitions:

$$(A \cup B)(x) = \min\left(A(x) + B(x), 1\right),$$
$$(A \cap B)(x) = \max\left(A(x) + B(x) - 1, 0\right),$$

for every $x \in X$. The union and intersection of fuzzy sets can be extended to the finite case in an obvious way (see, e.g., Bede (2013)). Because $A \cup B = X$ and $A \cap B = \varnothing$ if and only if $A(x) + B(x) = 1$, for every $x \in X$, a natural definition of a fuzzy partition can be given.

Definition 1. (see, e.g., Dumitrescu, Lazzerini, and Jain, 2000, p. 70) The family $\{A_1, \ldots, A_s\}$, $s > 2$, of fuzzy sets on X is called disjoint if and only if the equality

$$\left(\bigcup_{i=1}^{r} A_i\right) \cap A_{r+1} = \varnothing$$

holds for $r \in \{1, \ldots, s-1\}$.

Definition 2. (see, e.g., Dumitrescu et al., 2000, p. 72) Let $A_1, \ldots, A_s$, $s > 2$, be fuzzy sets on X. The family $P = \{A_1, \ldots, A_s\}$ is a fuzzy partition of X if and only if the following requirements are fulfilled:

(i) P is a disjoint family of fuzzy sets
(ii) $\cup_{i=1}^{s} A_i = X$.

An element $A_i \in P$ is called an atom of the fuzzy partition P. A characterization of a fuzzy partition is given by (see Dumitrescu et al., 2000, pp. 72–73):

Proposition 3. *$P = \{A_1, \ldots, A_s\}$ is a fuzzy partition of X if and only $\sum_{i=1}^{s} A_i(x) = 1$, for every $x \in X$.*

It is immediate that we may associate an $s \times n$ matrix $Q = (q_{ij})$, $i \in \{1, \ldots, s\}$, $j \in \{1, \ldots, n\}$ with every fuzzy partition $P = \{A_1, \ldots, A_s\}$ of a finite set $X = \{x^1, \ldots, x^n\}$, where $q_{ij} = A_i(x^j)$ is the membership degree of the element x^j to the atom $A_i \in P$, and $\sum_{i=1}^{s} q_{ij} = 1$, for every $j \in \{1, \ldots, n\}$. Q may be called the partition matrix or the matrix representation of the fuzzy partition P. If Q_i is the matrix representation of the fuzzy partition P^i, $i \in \{1, 2\}$ then we define the distance between P^1 and P^2 as (see Dumitrescu et al., 2000)

$$D(P^1, P^2) = \|Q_1 - Q_2\|,$$

where $\|\cdot\|$ may be defined in various ways. For example,

$$\|Q\| = \max_{i \in \{1,\ldots,s\}, j \in \{1,\ldots,n\}} |q_{ij}|.$$

It is well known that the fuzzy clustering is an important tool to identify the structure in data (see, e.g., Bezdek, 1981; Bezdek, Ehrlich, & Full, 1984; Cundari, Sârbu, & Pop, 2002; Dumitrescu et al., 2000; Kandel, 1982; Mei & Chen, 2014; Wang, Ma, Lao, & Wang, 2014; Zhao, Fan, & Liu, 2014) and, in addition, it does not require a very sophisticated background.

Let $X = \{x^1, \ldots, x^n\} \subset \mathbb{R}^p$ be a set of vectors, where n is the number of objects and p is the number of characteristics, $x^j = (x^j_1, \ldots, x^j_p)$, and $L = \{L^1, \ldots, L^s\}$ be a s-tuple of prototypes, $L^i = (L^i_1, \ldots, L^i_p)$, each of them characterizing one of the s clusters of the data set. A partition of X into s fuzzy clusters is performed by minimizing the objective function (Cundari et al., 2002; Dumitrescu et al., 2000; Dumitrescu, Sârbu, and Pop, 1994; Pop, Sârbu, Horowitz, and Dumitrescu, 1996)

$$J(P, L) = \sum_{i=1}^{s} \sum_{j=1}^{n} (A_i(x^j))^2 d^2(x^j, L^i), \tag{1}$$

where $P = \{A_1, \ldots, A_s\}$ is the fuzzy partition of X, $A_i(x^j) \in [0, 1]$ represents the membership degree of a point x^j to cluster A_i and d is a distance on $\mathbb{R}^p$, usually the Euclidean distance, that is

$$d^2(x^j, L^i) = \sum_{k=1}^{p} \left(x^j_k - L^i_k\right)^2. \tag{2}$$

For a given set of prototypes L, the minimum of the function $J(\cdot, L)$ is obtained for (see Dumitrescu et al., 2000)

$$A_i(x^j) = \frac{1}{\sum_{k=1}^{s} \frac{d^2(x^j, L^i)}{d^2(x^j, L^k)}}, \quad i \in \{1, \ldots, s\}, \quad j \in \{1, \ldots, n\}. \tag{3}$$

For a given partition P, the minimum of the function $J(P, \cdot)$ is obtained for (see Dumitrescu et al., 2000)

$$L^i = \frac{\sum_{j=1}^{n} (A_i(x_j))^2 x^j}{\sum_{j=1}^{n} (A_i(x_j))^2}, \quad i \in \{1, \ldots, s\}. \tag{4}$$

The optimal fuzzy partition of X is determined by using an iterative method, where J is successively minimized with respect to P and L. The procedure, described below, is called Fuzzy C-Means Algorithm (see Bezdek, 1981; Dumitrescu et al., 2000, pp. 293–295).