IIMB Management Review (2010) 22, 16—24

available at www.sciencedirect.com

ScienceDirect

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

Attribute selection in marketing: A rough set approach

Sabita Mahapatra [a], Sreekumar [b], S.S. Mahapatra [c,*]

[a] *Marketing, Indian Institute of Management Indore, Pigdamber, Rau, Indore 453331, Madhya Pradesh, India*
[b] *Rourkela Institute of Management Studies, Rourkela 769015, India*
[c] *Department of Mechanical Engineering, National Institute of Technology, Rourkela 769008, India*
Available online 13 April 2010

KEYWORDS
Rough set;
Marketing;
Attributes;
Rule discovery

Abstract Using an illustrative case study on the Indian cosmetic industry, this paper illustrates the advantages of the rough set approach over conventional techniques for the extraction of decision rules from data sets, which can be useful in various marketing applications. The rule generated through the methodology can act as an 'expert', which may be referred to in future strategic decision-making. The approach gives results similar to the results obtained through statistical methods but without making any assumption.

Introduction

The IT revolution has radically changed the way data is collected and generated, facilitating the process of decision making. A huge set of data has no practical relevance unless it can be mined to provide useful information pertaining to the interests of the organisation. The patterns in the data need to be deciphered in order to gain insights about aspects such as customer preferences, market trends, and business performance. Quick responses to the changing market environment are possible only if timely and accurate insights into the business and the market conditions are readily available (Sreekumar & Panda, 2005). The perpetually growing volume of data needs to be reduced into useful information, and this calls for tools that are capable of distinguishing the various properties of the data collected/generated. Such dimensionality reduction would enable companies to remain more focused, and would thereby reduce the labour and communication costs for data collection. Researchers generally tend to apply statistical inferences on the existing data, emphasising efficient use of organisational data through data mining and data warehousing (Ha & Park, 1998). However, the conventional techniques cannot reduce the data dimensions efficiently and the persistent redundant attributes would affect the rule discovery process, leading to highly degraded rules (Zhong, Dong, & Ohsugu, 2001).

All these factors have opened up the scope for some of the newer techniques which have been developed in recent years (Beynon, Curry, & Morgan, 2001). In this paper, the rough set theory (RST) developed by Pawlak (1982) is adopted as an alternative technique for the extraction of decision rules from data sets. The rough set approach has several advantages over the conventional methods (Dimitras, Slowinski, Susmaga, & Zopounidis, 1999; Shen & Loh,

* Corresponding author. Tel.: +91 0661 2462512; fax: +91 0661 2472926.
E-mail addresses: mahapatrass2003@yahoo.com, ssm@nitrkl.ac.in (S.S. Mahapatra).
 Peer-review under responsibility of Indian Institute of Management Bangalore.
doi:10.1016/j.iimb.2010.03.001

Production and hosting by Elsevier

2004). The tool is based on the original data, and does not need any external information. It is a tool suitable for analysing quantitative as well as qualitative attributes. This tool discovers important facts hidden in the data set, and expresses these facts in the natural language of decision rules or computational algorithms, and not as mathematical functional forms (Wolfram, 2002). The former is better at pattern recognition than the latter, and so can have better managerial applications. The set of derived decision rules gives a generalised description of the knowledge contained in the database, eliminating any redundancy inherent in the original data. The derived decision rules are based on facts—each decision rule is supported by a set of real examples. The results of rough sets are easy to understand and process, while the results of the other methods usually require an interpretation of the technical parameters with which the user may not be familiar.

This paper uses the basic ideas of RST to show how rule discovery can be made, and to present the relationship between the attributes. The fundamental concepts of the rough set approach are briefly explained in the following section. A case study is undertaken based on the data collected from twenty-three Indian cosmetic companies. The names of the companies have been withheld to maintain confidentiality.

Fundamentals of rough set theory

The rough set theory (RST) was developed by Pawlak (1982) at the Institute of Computer Sciences, Warsaw. It was initially proposed as an alternative data analysis method but subsequently found application in the areas of artificial intelligence, knowledge discovery, decision analysis, and expert systems among others. RST can deal with inexact, uncertain, and vague datasets (Shyng, Wang, Tzeng, & Wu, 2007); it is a new mathematical approach to vagueness. According to Pawlak and Skowron (2007), the rough set philosophy was founded on the assumption that some information (data, knowledge) is associated with every object of the universe of discourse. For example, if the objects under study are patients suffering from a certain disease, the symptoms of the disease form the information about the patients. Objects characterised by the same information are indiscernible (similar) in view of the available information about them. The indiscernibility relation generated in this way is the mathematical basis of RST. The rough set theory has been applied in various fields like marketing, banking, engineering, and medicine among others.

In RST, data is represented through a data table also known as an attribute-value table, an information table or a database. The rows of the table stand for the *objects*, the columns represent the *attributes*, and the entries are called *attribute values*. A database S is a pair represented by $S = \{U, A\}$, where U and A are both finite non empty sets; U is the universal set and A is the set of attributes. The subset of attributes in the database is the cluster of objects having the same attribute values or the same features. Objects that have the same features are indiscernible (similar); these blocks provide the elementary granules of knowledge. These granules are called concepts or elementary sets, and form the elementary building blocks (atoms) of knowledge. Any union of elementary sets is called a crisp (precise) set, and any other set is referred to as a rough set (vague, imprecise). Associated with every set X, there are two crisp sets called the lower and the upper approximations of X. The lower approximation of X is the union of all the elementary sets which are included in X, and the upper approximation of X is the union of all the elementary sets which have a non-empty intersection with X.

Exhibit 1 illustrates these concepts using the data of six objects (companies), the three attributes $\{a_1, a_2, a_3\}$, and the decision state D (Profit); the three attributes used are availability of research and development (R&D) facilities, adoption of state of the art technology, and marketing expenditure. These three variables could be considered to represent the independent variable, and proper utilisation of these resources in the right combination could lead the company to profit. Profit is the result of the decision variable, and it can take two attribute values—yes or no; so Profit can be considered to be the decision state. In principle, there can be more than one decision variable, but in this case only the decision variable Profit has been considered, which is a *distinguished* attribute. The methodology demonstrated here could also be used to discover rules that distinguish small profit-making firms from large loss-making firms.

$C3$ has R&D facilities and state of the art technology, and spends a very high amount for its marketing activities, and the attribute value of the decision variable (Profit) is Yes; i e, the company $C3$ makes profits. The information about $C3$ has the following attribute values: (a_1, Yes), (a_2, *Yes*), (a_3, *Very High*), (D, *Yes*).

Let U denote the set of all cases, A the set of all attributes, and V the set of all attribute values. A table such as the one in Exhibit 1 would define an information function ρ: $U \times A \rightarrow V$. The attribute values can also be written as a function, of the form $\rho(C3, a_1) = Yes$.

Let $a \in A$, $v \in V$, and $t = (a, v)$ be an attribute-value pair. A block of t, denoted by $[t]$, is a set of all cases from U for which the attribute a has the value v. So the information

Exhibit 1 Representation of sample database with six objects and three attributes.

Company	Availability of R&D facilities (a_1)	Adoption of state of the art technology (a_2)	Marketing expenditure (a_3)	Profit (D)
$C1$	No	Yes	High	Yes
$C2$	Yes	No	High	Yes
$C3$	Yes	Yes	Very high	Yes
$C4$	No	Yes	Average	No
$C5$	Yes	No	High	No
$C6$	No	Yes	Very high	Yes