

# Chapter 3

## Data Mining Using Fuzzy Decision Trees:

### An Exposition from a Study of Public Services Strategy in the USA

**Malcolm J. Beynon**  
Cardiff University, UK

**Martin Kitchener**  
Cardiff Business School, UK

#### ABSTRACT

*The chapter expositis the strategies employed by the public long-term care systems operated by each U.S. state government. The central technique employed in this investigation is fuzzy decision trees (FDTs), producing a rule-based classification system using the well known soft computing methodology of fuzzy set theory. It is a timely exposition, with the employment of set-theoretic approaches to organizational configurations, including the fuzzy set representation, starting to be discussed. The survey details considered, asked respondents to assign each state system to one of the three 'orientations to innovation' contained within Miles and Snows' (1978) classic typology of organizational strategies. The instigated aggregation of the experts' opinions adheres to the fact that each long-term care system, like all organizations, is "likely to be part prospector, part defender, and part reactor, reflecting the complexity of organizational strategy". The use of FDTs in the considered organization research problem is pertinent since the linguistic based fuzzy decision rules constructed, open up the ability to understand the relationship between a state's attributes and their predicted position in a general strategy domain - the essence of data mining.*

#### INTRODUCTION

With data storage increasing at a phenomenal rate, traditional *ad hoc* mixtures of data mining tools are no longer adequate. In one response, some attention

has been given to the potential for soft computing frameworks to provide flexible information processing capability that can exploit the tolerance of imprecision, uncertainty, approximate reasoning, and partial truth in knowledge discovery (Mitra *et al.*, 2002). This chapter extends that line of enquiry

DOI: 10.4018/978-1-60566-906-9.ch003

by providing an early and detailed exposition of the data mining potential of a soft computing methodology that is based on fuzzy set theory, henceforth FST (Zadeh, 1965).

Since its introduction in 1965, FST is closely associated with uncertain reasoning and is the earliest and most widely reported constituent of soft computing (Mitra *et al.*, 2002). Of particular interest in this exposition of data concerning public policy and strategy, FST incorporates opportunities to develop techniques that incorporate vagueness and ambiguity in their operation, and it allows outputs to be presented in a highly readable and easily interpretable manner (Zhou and Gan, 2008). While data mining encompasses the typical tasks of; classification, clustering, association and outlier detection, here its role in rule-based classification is considered.

Previous FST-based research in organizational and policy contexts is limited but includes: explaining constitutional control of the executive of parliamentary democracies in US states (Pennings, 2003), and the evaluation of knowledge management capability of organizations (Fan *et al.*, 2009). Ragin and Pennings (2005) give a discussion of FST in social research, in their introduction to a special issue of the journal *Sociological Methods & Research*. This acknowledges the need to continually validate this new methodology (FST), through its continued application. A pertinent study by Fiss (2007), considered the whole issue of the employment of a set-theoretic approach to organizational configurations, including the progression from a crisp to fuzzy set representation, and the latter's potential for undertaking appropriate analysis.

The context of the exposition presented in this chapter is a study of the strategies employed by the public long-term care systems operated by each U.S. state government. The main dataset was collected from a survey of experts in this area (including academics, government officials, and service providers). The survey asked respondents to assign each state system to one of the three 'orientations to innovation' contained within

Miles and Snows' (1978) classic typology of organizational strategies: prospectors, defenders, and reactors. Briefly, these strategic groups describe different orientations to strategy from the more consistently innovative prospectors, to reactors that typically innovate only after coercion. The instigated aggregation of the experts' opinions adheres to the fact that each long-term care system, like all organizations, is "likely to be part prospector, part defender, and part reactor, reflecting the complexity of organizational strategy" (Andrews *et al.*, 2006).

In this chapter, the aggregated expert assignments are assessed using a fuzzy decision tree (FDT) analysis of state long-term care system characteristics. The pertinence of this analysis is that, with the federal system existing in each U.S. state, the decision rules constructed are in respect of the state's governing organization's management attitudes to healthcare. FDT is a data mining technique which benefits from the general methodology FST (Yuan and Shaw, 1995; Mitra *et al.*, 2002). The overriding remit of, decision trees, within crisp and fuzzy environments, is with the classification of objects described by a data set in the form of a number of condition and decision attributes.

A decision tree, in general, starts with an identified root node, and paths are constructed down to leaf nodes, where the attributes associated with the intermediate nodes are identified through a measure to preferentially gauge the classification certainty of certain objects down that path. Each path down to a leaf node forms an 'if.. then..' decision rule, used to classify those objects whose condition attribute values satisfy the condition part of that rule. Beyond FDT, other rule based classification methods include, amongst others, RIPPER (Cohen, 1995; Thabtah *et al.*, 2006) and rough set theory (Beynon *et al.*, 2000).

The development of decision trees in a fuzzy environment furthered the readability of the now constructed 'if.. then..' fuzzy decision rules (Zhou and Gan, 2008). The potential appropriateness of

18 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/data-mining-using-fuzzy-decision/44282](http://www.igi-global.com/chapter/data-mining-using-fuzzy-decision/44282)

## Related Content

---

### Advances in Classification of Sequence Data

Pradeep Kumar, P. Radha Krishna, Raju S. Bapi and T. M. Padmaja (2008). *Data Mining and Knowledge Discovery Technologies* (pp. 143-174).

[www.irma-international.org/chapter/advances-classification-sequence-data/7517](http://www.irma-international.org/chapter/advances-classification-sequence-data/7517)

### Combining BPSO and ELM Models for Inferring Novel lncRNA-Disease Associations

Wenqing Yang, Xianghan Zheng, QiongXia Huang, Yu Liu, Yimi Chen and ZhiGang Song (2023). *International Journal of Data Warehousing and Mining* (pp. 1-18).

[www.irma-international.org/article/combining-bpso-and-elm-models-for-inferring-novel-lncrna-disease-associations/317092](http://www.irma-international.org/article/combining-bpso-and-elm-models-for-inferring-novel-lncrna-disease-associations/317092)

### Spatio-Temporal OLAP Queries Similarity Measure and Algorithm

Olfa Layouni and Jalel Akaichi (2019). *International Journal of Data Warehousing and Mining* (pp. 22-41).

[www.irma-international.org/article/spatio-temporal-olap-queries-similarity-measure-and-algorithm/225805](http://www.irma-international.org/article/spatio-temporal-olap-queries-similarity-measure-and-algorithm/225805)

### Data Quality for Data Mining in Business Intelligence Applications: Current State and Research Directions

Arun Thotapalli Sundararaman (2015). *Integration of Data Mining in Business Intelligence Systems* (pp. 34-59).

[www.irma-international.org/chapter/data-quality-for-data-mining-in-business-intelligence-applications/116806](http://www.irma-international.org/chapter/data-quality-for-data-mining-in-business-intelligence-applications/116806)

### Web Usage Mining and the Challenge of Big Data: A Review of Emerging Tools and Techniques

Abubakr Gafar Abdalla, Tarig Mohamed Ahmed and Mohamed Elhassan Seliaman (2016). *Big Data: Concepts, Methodologies, Tools, and Applications* (pp. 899-928).

[www.irma-international.org/chapter/web-usage-mining-and-the-challenge-of-big-data/150199](http://www.irma-international.org/chapter/web-usage-mining-and-the-challenge-of-big-data/150199)