Chapter 6
Clustering Approaches in Decision Making Using Fuzzy and Rough Sets

Deepthi P. Hudedagaddi
VIT University, India

B. K. Tripathy
VIT University, India

ABSTRACT

Data clustering has been an integral and important part of data mining. It has wide applications in database anonymization, decision making, image processing and pattern recognition, medical diagnosis and geographical information systems, only to name a few. Data in real life scenario are having imprecision inherent in them. So, early crisp clustering techniques are very less efficient. Several imprecision based models have been proposed over the years like the fuzzy sets, rough sets, intuitionistic fuzzy sets and many of their generalized versions. Of late, it has been established that the hybrid models obtained as combination of these imprecise models are far more efficient than the individual ones. So, many clustering algorithms have been put forth using these hybrid models. The focus of this chapter is to discuss on some of the data clustering algorithms developed so far and their applications mainly in the area of decision making.

On an important decision one rarely has 100% of the information needed for a good decision no matter how much one spends or how long one waits. And, if one waits too long, he has a different problem and has to start all over. This is the terrible dilemma of the hesitant decision maker. -Robert K. Greenleaf

INTRODUCTION

A cluster is a collection of data elements that are similar to each other but dissimilar to elements in other clusters. A vast amount of data is generated and made available across multiple sources. It is practically impossible to manually analyze the myriad of data and select the data that is required to perform a particular task. Hence, a mechanism that can classify the data according to some criteria in which only the classes of interest are selected and rests are rejected is essential. Clustering techniques are applied in the analysis of statistical data used in fields such as machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics and is a major task in exploratory data mining (Bezdek & Pal, 1998; Tou & Gonzalez, 1974). A wide number of clustering algorithms have been proposed to suit the requirements in each field of its application.

SETS

The notion of a set is not only basic for the whole mathematics but it also plays an important role in natural language. Sets include collections of various objects of interest, e.g., collection of books, paintings, people etc. The notion of set was proposed by G. Cantor in 1983. It was not free from controversies like that of B. Russel which questioned the definition of a set proposed by Cantor as a collection of distinct and distinguishable elements. Intuitively a set can be defined as a well- defined collection of objects called elements.

Rough Sets

A rough set, first described by a Polish computer scientist Z. Pawlak (1982), is a formal approximation of a crisp set (i.e., conventional set) in terms of a pair of sets which give the lower and the upper approximation of the original set. In the standard version of rough set theory, the lower- and upper-approximation sets are crisp sets, but in other variations, the approximating sets may be fuzzy sets.

Definition of a Rough Set

A set X in a universal set U is considered to be rough or not with respect to an equivalence relation P defined over U. Since P is an equivalence relation, it decomposes U into nonempty disjoint equivalence classes. These equivalence classes are called the categories and are the granules of knowledge induced by P over U. The equivalence class of an element x with respect to P is denoted by $[x]_P$. By U/P, the set of equivalence classes generated by P over U is represented.

Pair of crisp sets is associated with any set X in U with respect to P and are called the P-lower approximation of X and P-upper approximation of X. They are denoted by $\underline{P}X$ and $\overline{P}X$ respectively, and defined as

$$\underline{P}X = \{x \mid [x]_P \subseteq X\} \tag{1}$$

$$\overline{P}X = \{x \mid [x]_P \bigcap X \neq \phi\} \tag{2}$$

Related Content

Developing Fuzzy-AHP-Integrated Hybrid MCDM System of COPRAS-ARAS for Solving an Industrial Robot Selection Problem
Shankha Shubhra Goswamiand Dhiren Kumar Behera (2023). *International Journal of Decision Support System Technology (pp. 1-38).*
www.irma-international.org/article/developing-fuzzy-ahp-integrated-hybrid-mcdm-system-of-copras-aras-for-solving-an-industrial-robot-selection-problem/324599

Introducing a Conceptual Framework for Spatial Decision SIM: A Tool Supports Decisions for Stakeholders
Ahmad M. A. Toimahand Samy M. Z. Afifi (2021). *Research Anthology on Decision Support Systems and Decision Management in Healthcare, Business, and Engineering (pp. 796-813).*
www.irma-international.org/chapter/introducing-a-conceptual-framework-for-spatial-decision-sim/282616

Backward and Forward Linkages in Chinese Steel Industry Using Input Output Analysis
Lafang Wang, Rui Xieand Jun Liu (2013). *Management Theories and Strategic Practices for Decision Making (pp. 139-158).*
www.irma-international.org/chapter/backward-forward-linkages-chinese-steel/70955

Towards Identifying Multicriteria Outliers: An Outranking Relation-Based Approach
Baroudi Roubaand Safia Nait-Bahloul (2018). *International Journal of Decision Support System Technology (pp. 27-38).*
www.irma-international.org/article/towards-identifying-multicriteria-outliers/205660

Evaluation of a Multi-Goal Solver for Use in a Blackboard Architecture
Jeremy Straub (2014). *International Journal of Decision Support System Technology (pp. 1-13).*
www.irma-international.org/article/evaluation-of-a-multi-goal-solver-for-use-in-a-blackboard-architecture/117685