LEARNING DECISION RULES FROM UNCERTAIN DATA USING ROUGH SETS

Salsabil TRABELSI and Zied ELOUEDI

Institut Superieur de gestion, larodec laboratory
41 Avenue liberte, Bardo, Tunisie
**E-mail: salsabilt@yahoo.fr*
E-mail: Zied.elouedi@gmx.fr

In this paper, we deal with the problem of learning decision rules from partially uncertain data based on rough sets. The uncertainty exists in the decision attribute and not in condition attribute values of the decision system. This latter is represented by the belief function theory. So, we will adapt the basic concepts of rough sets in order to generate rules, denoted belief decision rules.

Keywords: uncertainty, belief function theory, rough sets, classification.

1. Introduction

The rough set theory based on approximation reasoning proposed by Pawlak 4, 5] constitutes a sound basis for data mining. It allows reducing original data and generating in automatic way the sets of decision rules from data (rough set classification). One enhancement of standard rough set classification does not well perform their task in an environment characterized by uncertainty or incomplete data. Many researches have been done to adapt rough sets to this kind of environment 1, 2, 6, 11]. These extensions do not deal with partially uncertain decision attribute values in decision system. This kind of uncertainty exists in many real-world applications like in medicine where diseases of some patients may be partially uncertain. In this paper, we will propose a new approach of learning decision rules from partially uncertain data using rough sets. This uncertainty exists in decision attribute values of decision system. The uncertainty is represented under the belief function framework which is able to handle the partial or total ignorance in a flexible way. The relationship between rough sets and belief functions continues to receive the attention of many authors 3, 10]. This paper is organized as follows: Section 2 provides an overview about the rough set theory. Section 3 introduces the belief function theory as understood in the transferable belief model (TBM). Section 4 describes our new approach of learning decision rules by rough sets to handle the problem of uncertainty under the belief function framework.

2. Rough set theory

This section presents the basic concepts of rough sets proposed by Pawlak 4, 5] to deal with vague concepts. $A = (\mathrm{U},\ \mathrm{C} \cup \{\mathrm{D}\})$ is a decision system, where U is a finite set of objects and C is finite set of *condition* attributes, i.e., c:U $\rightarrow Vc$ for c $\in C$ (Vc is called the value set of attribute c). In supervised learning, $D \neq C$ is a distinguished attribute called *decision*. The rough sets adopts the concepts of indiscernibility relation to partition training instances according to some criteria. The objects x and x' are indiscernible on a subset of attributes B, if they have the same values for each attribute in subset B of A. The equivalence classes thus partitions the object set U into disjoint subsets, denoted by U/B, and the partition including x is denoted $[x]_B$. The rough set approach analyses data according to two basic concepts the lower and upper approximations for B on X, denoted $\underline{\mathrm{B}}$X and $\bar{\mathrm{B}}$X respectively where

$$\underline{\mathrm{B}}\mathrm{X} = \{\mathrm{x} \mid [x]_B \subseteq X\ \} \quad \text{and} \quad \bar{\mathrm{B}}\mathrm{X} = \{\mathrm{x} \mid [x]_B \cap X \neq \emptyset\}$$

The objects in $\underline{\mathrm{B}}$X can be certainty classified as members of X, while the objects in $\bar{\mathrm{B}}$X can be only classified as possible members of X. After the lower and the upper approximations have been found, the rough set theory can be then used to derive certain and possible rules from them.

3. Belief function theory

In this section, we briefly review the main concepts underlying the TBM, one interpretation of the belief function theory 13]. Let Θ be a finite set of elementary events to a given problem, called the frame of discernment 12]. All the subsets of Θ belong to the power set of Θ, denoted by 2^Θ. The impact of a piece of evidence on the different subsets of the frame of discernment Θ is represented by a basic belief assignment (bba). The bba is a function $m : 2^\Theta \rightarrow [0, 1]$ such that:

$$\sum_{A \subseteq \Theta} m(A) = 1 \tag{1}$$

The value *m(A)*, named a basic belief mass (bbm), represents the portion of belief committed exactly to the event A. In the transferable belief model, holding beliefs and making decisions are distinct processes. Hence, it proposes two level models:

- *The credal level* where beliefs are represented by belief functions.
- *The pignistic level* where beliefs are used to make decisions and represented by probability functions called the pigninstic probabilities and is defined as: 13]

$$BetP(A) = \sum_{B \subseteq \Theta} \frac{| A \cap B |}{| B |} \frac{m(B)}{(1 - m(\emptyset))}, \quad \text{for all} \quad A \in \Theta \qquad (2)$$

4. Rough sets under uncertainty

In this Section, a new version of learning decision rules using rough sets from partially uncertain decision system is proposed. The uncertainty is represented by the TBM and exists only in decision attribute values. So, we will adapt the basic concepts of rough sets to deal with this situation.

4.1. *Decision system under uncertainty*

Our uncertain decision system denoted U contains n objects O_j, characterized by m certain condition attribute and uncertain decision attribute. We propose to represent the uncertainty of each object by a bba $m^{\Theta}\{O_j\}$ expessing belief on classes defined on the frame of discernment Θ.

Example: Let us take Table 1 to describe our uncertain decision system. This latter contains eight objects, three certain condition attributes C={Headache, Muscle-pain, Temperature} and an uncertain decision attribute UD={Flu} with possible value {yes, no} representing Θ.

Table 1. Uncertain decision table.

Patient	Headache	Muscule-pain	Temperature	Flu
O_1	yes	yes	very high	$m^{\Theta}\{O_1\}(yes) = 1$
O_2	yes	no	high	$m^{\Theta}\{O_2\}(yes) = 1$
O_3	yes	yes	high	$m^{\Theta}\{O_3\}(yes) = 0.5 \quad m^{\Theta}\{O_3\}(\Theta) = 0.5$
O_4	no	yes	normal	$m^{\Theta}\{O_4\}(no) = 0.6 \quad m^{\Theta}\{O_4\}(\Theta) = 0.4$
O_5	no	yes	normal	$\mathrm{m}^{\Theta}\{O_5\}(no) = 1$
O_6	yes	no	high	$\mathrm{m}^{\Theta}\{O_6\}(no) = 1$
O_7	no	yes	very high	$\mathrm{m}^{\Theta}\{O_7\}(yes) = 1$
O_8	no	no	normal	$\mathrm{m}^{\Theta}\{O_8\}(no) = 1$

Note: For the patient O_3, 0.5 of beliefs are exactly committed to the decision d_1=yes, whereas 0.5 of beliefs is assigned to the whole of frame of discernment Θ (ignorance).

4.2. *Indiscernibility relation*

Indiscernibility relation for the condition attributes is the same as in the certain case because their values are certain. In our case is equal to: U/C={$\{O_1\}$, $\{O_2, O_6\}$, $\{O_3\}$, $\{O_4, O_5\}$, $\{O_7\}$, $\{O_8\}$}. Indiscernibility relation for the decision attribute is not the same as in the certain case. The decision value is represented by a bba. In our case, we have two equivalence classes d_1=yes and d_2=no. So, we need for optimal decision making to assign each object to the most probable class. The idea is to use the pignistic transformation. It is a function which can transform the belief function to probability function in order to make decisions from beliefs. We suggest, for each object O_j in the decision system U, compute the pignistic probability, denoted BetP, by applying the pignistic transformation to $m^{\Theta}\{O_j\}$.
Example: Let continue with the same example. Table 2 shows the pignistic probability applying to each $m^{\Theta}\{O_j\}$.

Table 2. Pignistic transformation.

m^{Θ}	BetP
$m^{\Theta}\{O_1\}$	BetP(yes)=1 BetP(no)=0
$m^{\Theta}\{O_2\}$	BetP(yes)=1 BetP(no)=0
$m^{\Theta}\{O_3\}$	BetP(yes)=0.75 BetP(no)=0.25
$m^{\Theta}\{O_4\}$	BetP(yes)=0.2 BetP(no)=0.8
$m^{\Theta}\{O_5\}$	BetP(yes)=0 BetP(no)=1
$m^{\Theta}\{O_6\}$	BetP(yes)=0 BetP(no)=1
$m^{\Theta}\{O_7\}$	BetP(yes)=1 BetP(no)=0
$m^{\Theta}\{O_8\}$	BetP(yes)=0 BetP(no)=1

The object O_3 included in d_1=yes, because it is the most probable class. So, the two equivalence classes based on uncertain decision attribute are as follows: U/UD={$\{O_1,O_2,O_3,O_7\}$, $\{O_4,O_5,O_6,O_8\}$}. If we have equal probabilities, we choose one of decision attribute values arbitrarily.

4.3. *Set approximation*

To compute the new lower and upper approximation for our uncertain decision table, we follow two steps:

(1) For each equivalence classes based on condition attributes C, combine their bba using the operator mean. In order to check which of them has certain bba.

(2) For each equivalence classes X based on uncertain decision attribute, we compute the new lower and upper approximation, as follows:

$$\underline{C}X=\{x|\ [x]_C \subseteq X \text{ and } m^{\Theta}(d)=1\} \text{ and } \bar{C}X=\{x|\ [x]_C \cap X \neq \emptyset\ \}$$

Example: We continue with the same example to compute the new lower and upper approximation. After the first step, we obtain the combined bba for each equivalence classes U/C using operator mean. Table 3 represents the combined bba for the subset $\{O_4,\ O_5\}$.

Table 3. combined bba for $\{O_4,\ O_5\}$.

Patient	$m^{\Theta}(yes)$	$m^{\Theta}(no)$	$m^{\Theta}(\Theta)$
O_4	0	0.4	0.6
O_5	0	1	0
m	0	0.7	0.3

Next, we compute the lower and upper approximation for each equivalence classes U/UD. For d_1=yes, let X= $\{O_1,O_2,O_3,O_7\}$

$$\underline{C}X=\{\{O_1\},\ \{O_7\}\} \text{ and } \bar{C}X=\{\{O_1\},\ \{O_2,\ O_6\},\ \{O_3\},\ \{O_7\}\}$$

For d_2=no, let Y= $\{O_4,O_5,O_6,O_8\}$

For example, the subset $\{O_4,\ O_5\}$ is included to Y and it has uncertain bba. So, we put it in the upper.

$$\underline{C}Y=\{\{O_8\}\} \text{ and } \bar{C}Y=\{\{O_2,\ O_6\},\ \{O_4,\ O_5\},\ \{O_8\}\}$$

We can generate the decision rules from our uncertain decision table. One of certain rules induced from lower approximation like in $\{O_1\}$:
"*If* Headache=yes *and* Muscle-pain=yes *and* Temperature= very high *then* Flu=yes".
One of uncertain rules induced from upper approximation like in $\{O_3\}$:
"*If* Headache=yes *and* Muscle-pain=yes *and* Temperature= high *then* $m^{\Theta}(yes)=0.5$ $m^{\Theta}(\Theta)=0.5$. It can be denoted belief decision rule.

5. Conclusion and future work

In this paper, we have proposed a new learning approach to derive decision rules from uncertain decision system using rough sets. The uncertainty in the decision system exists only in decision attribute values. This latter is represented by the belief function theory as understood by the TBM. The belief decision rules generated from our decision table are not optimal. As a future work, we suggest improve our new approach of learning belief decision rules based on rough sets. We try to propose algorithm of simplification decision table and generation algorithm of significant rules.

References

1. Grzymala-Busse, J.: Modified algorithms LEM1 and LEM2 for rule induction from data with missing attribute values. Proc of the fifth international workshop on Rough sets and soft computing at the third joint conference on Information Sciences, (1997) 69-72.
2. Kryszkiewiez, M.: Rough set approach to incomplete information systems, Proceedings of the second annual joint conference on Information Science, Wrightsville, NC (1995) 194-197.
3. Lingras, P.J.: Combination of Evidence in Rough Set Theory, Proceedings of the Fifth International Conference on Computing and Information, Sudbury, Canada (1993) 289-293.
4. Pawlak, Z.: Rough Sets, International Journal of Computer and Information Sciences, Vol 11, (1982) 341-356.
5. Pawlak, Z., Zdzislaw, A.: Rough Sets: Theoretical Aspects of Reasoning About Data. Dordrecht: Kluwer Academic Publishing. ISBN 0-7923-1472-7(1991).
6. Orlowska, E.(ed.) Incomplete Information: Rough Set Analysis, Physical-Verlag (1997).
7. Shafer, G.: A mathematical theory of evidence. Princeton University Press. Princeton, NJ (1976).
8. Skowron, A. et all: Rough set perspective on data and knowledge, Handbook of Data Mining and Knowledge Discovery, Oxford University Press (2002) 134-149.
9. Skowron, A.: Synthesis of adaptive decision systems from experimental data. Proc. Fifth Scandinavian Conference on Artificial Intelligence. Trondheim, Norway (1995) 220-238.
10. Skowron, A., Grzymala-Busse, J.: From rough set theory to evidence theory. New York(1994) 193-236.
11. Slowinski, R., Stefanowski, J.: Rough classification in incomplete information systems, Mathematical and Computer Modeling, Vol 12, (1989) 1347-1357.
12. Smets, P.: Belief functions. In P. Smets, E. H. Mamdani, D. Dubois, H. Prade (Eds.): Non standard logics for automated reasoning. Academic Press, London (1988) 253-286.
13. Smets, P.: The transferable belief model for quantified belief representation. In D.M. Gabbay & P. Smets (Eds), Handbook of defeasible reasoning and uncertainty management systems, Vol 1, Doordrecht, The Netherlands: Kluwer (1998) 207-301.