Belief Rough Set Classifier

Salsabil Trabelsi¹, Zied Elouedi¹, and Pawan Lingras²

¹ Larodec, Institut Supérieur de Gestion de Tunis, Tunisia
² Saint Mary's University Halifax, Canada

Abstract. In this paper, we propose a new rough set classifier induced from partially uncertain decision system. The proposed classifier aims at simplifying the uncertain decision system and generating more significant belief decision rules for classification process. The uncertainty is repersented by the belief functions and exists only in the decision attribute and not in condition attribute values.

Keywords: rough sets, belief functions, uncertainty, classification.

1 Introduction

The standard rough set classifiers do not perform well in the face of uncertainty or incomplete data [3]. In order to overcome this limitation, researchers have adapted rough sets to the uncertain and incomplete data [2,5]. These extensions do not deal with partially uncertain decision attribute values in a decision system. In this paper, we propose a new rough set classifier that is able to learn decision rules from partially uncertain data. We assume that this uncertainty exists in decision attribute and not in condition attributes (we handle only symbolic attributes). We choose the belief function theory introduced by Shafer [4] to represent uncertainty which enables flexible representation of partial or total ignorance. The belief function theory used in this paper is based on the transferable belief model (TBM). The paper also provides experimental evaluation of our classifier on different databases.

2 Rough Sets

This section presents the basic concepts of rough sets proposed by Pawlak [3]. $A = (U, C \cup \{d\})$ is a decision system, where U is a finite set of objects $U = \{o_1, o_2, \ldots, o_n\}$ and C is finite set of *condition* attributes, $C = \{c_1, c_2, \ldots, c_s\}$. In supervised learning, $d \notin C$ is a distinguished attribute called *decision*. The value set of d, called $\Theta = \{d_1, d_2, \ldots, d_t\}$. For every object $o_j \in U$, we will use $c_i(o_j)$ to denote value of a condition attribute c_i for object o_j . Similarly, $d(o_j)$ is the value of the decision attribute for an object o_j . We further extend these notations for a set of attributes $B \subseteq C$, by defining $B(o_j)$ to be value tuple of attributes in B for an object o_j . The rough sets adopt the concepts of indiscernibility relation to

Y. Gao and N. Japkowicz (Eds.): Canadian AI 2009, LNAI 5549, pp. 257–261, 2009.

[©] Springer-Verlag Berlin Heidelberg 2009

258 S. Trabelsi, Z. Elouedi, and P. Lingras

partition the object set U into disjoint subsets, denoted by U/B, and the class that includes o_j is denoted $[o_j]_B$.

$$IND_B = U/B = \{ [o_j]_B \mid o_j \in U \}, \text{ with } [o_j]_B = \{ o_i \mid B(o_i) = B(o_j) \}$$
(1)

The equivalence classes based on the decision attribute is denoted by $U/\{d\}$

$$IND_{\{d\}} = U/\{d\} = \{[o_j]_{\{d\}} \mid o_j \in U\}$$
(2)

The lower and upper approximations for B on X, denoted $\overline{B}X$ and $\overline{B}X$ respectively where: $\overline{B}X = \{o_j \mid [o_j]_B \subseteq X\}$, and $\overline{B}X = \{o_j \mid [o_j]_B \cap X \neq \emptyset\}$

 $Pos_C(\{d\})$, called a positive region of the partition $U/\{d\}$ with respect to C.

$$Pos_C(\{d\}) = \bigcup_{X \in U/\{d\}} \mathbb{C}X$$
(3)

A reduct of C is a minimal set of attributes $B \subseteq C$ where:

$$Pos_B(\{d\}) = Pos_C(\{d\}) \tag{4}$$

3 Belief Function Theory

In this section, we briefly review the main concepts underlying the belief function theory as interpreted in the TBM [6]. Let Θ be a finite set of elementary events to a given problem, called the frame of discernment. 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 Θ is represented by a basic belief assignment (BBA). The BBA is a function $m: 2^{\Theta} \to [0, 1]$ such that: $\sum_{E \subseteq \Theta} m(E) = 1$.

In the TBM, the BBA induced from distinct pieces of evidence are combined by the rule of combination [6].

$$(m_1 \bigcirc m_2)(E) = \sum_{F,G \subseteq \Theta: F \cap G = E} m_1(F) \times m_2(G)$$
(5)

In the TBM, beliefs to make decisions can be represented by probability functions called the pignistic probabilities denoted BetP and is defined as [6]:

$$BetP(\{a\}) = \sum_{\emptyset \subset F \subseteq \Theta} \frac{|\{a\} \cap F|}{|F|} \frac{m(F)}{(1 - m(\emptyset))}, \text{ for all } a \in \Theta$$
(6)

4 Belief Rough Set Classifier

In this section, we will begin by describing the modified definitions of the basic concepts of rough sets under uncertainty. These adaptations were proposed originally in [7]. Second, we simplify the uncertain decision system to generate the more signification decision rules to create the belief rough sets classifier.

The basic concepts of rough sets under uncertainty are described as follows:

- 1. Decision system under uncertainty: Defined as follows: $A = (U, C \cup \{ud\})$, where U is a finite set of objects and C is finite set of *certain* condition attributes. $ud \notin C$ is an uncertain decision attribute. We propose to represent the uncertainty of each object by a BBA m_j expressing belief on decision defined on Θ representing the possible values of ud.
- 2. Indiscernibility relation: The indiscernibility relation for the decision attribute is not the same as in the certain case. The decision value is represented by a BBA. Therefore, we will use the $U//\{ud\}$ to denote the uncertain indiscernibility relation. We need to assign each object to the right decision classes. The idea is to use the pignistic transformation. For each object o_j in the decision system U, compute the pignistic probability, denoted $BetP_j$ to correspond to m_j . For every ud_i , a decision value, we define decision classes:

$$IND_{\{ud\}} = U//\{ud\} = \{X_i \mid ud_i \in \Theta\}, \text{ with } X_i = \{o_j \mid BetP_j(\{ud_i\}) > 0\}$$
(7)

- 3. Set approximation: To compute the new lower and upper approximations for our uncertain decision table, we follow two steps:
 - (a) For each equivalence class based on condition attributes C, combine their BBA using the operator mean. The operator mean is more suitable than the rule of combination [6] to combine these BBAs, because they are beliefs on decision for different objects and not different beliefs on one object. After combination, check which decision classes have certain BBA.
 - (b) For each decision class X_i based on uncertain decision attribute, we compute the new lower and upper approximations, as follows: $\underline{C}X_i = \{o_j \mid [o_j]_C \subseteq X_i \text{ and } m_j(\{ud_i\}) = 1\}$ and $\overline{C}X_i = \{o_j \mid [o_j]_C \cap X_i \neq \emptyset\}$
- 4. Positive region: We define the new positive region denoted $UPos_C(\{ud\})$.

$$UPos_C(\{ud\}) = \bigcup_{X_i \in U/ud} \underline{C}X_i,$$
(8)

5. **Reduct:** is a minimal set of attributes $B \subseteq C$ such that:

$$UPos_B(\{ud\}) = UPos_C(\{ud\}).$$
(9)

The main steps to simplify the uncertain decision system are as follows:

- 1. Eliminate the superfluous condition attributes: We remove the superfluous condition attributes that are not in reduct.
- 2. Eliminate the redundant objects: After removing the superfluous condition attributes, we will find redundant objects. They may not have the same BBA on decision attribute. So, we use their combined BBAs using the operator mean.

260 S. Trabelsi, Z. Elouedi, and P. Lingras

3. Eliminate the superfluous condition attribute values: In this step, we compute the reduct value for each belief decision rule R_j of the form: If $C(o_j)$ then m_j . For all $B \subset C$, let $X = \{o_k \mid B(o_j) = B(o_k)\}$ If $Max(dist(m_j, m_k)) \leq$ threshold then B is a reduct value of R_j . Where dist is a distance measure between two BBAs. We will choose the distance measure described in [1] which satisfies properties such as non-degeneracy and symmetry.

5 Experimental Results

We have performed several tests on real databases obtained from the U.C.I. repository ¹. These databases are modified in order to include uncertainty in the decision attribute. We use three degrees of uncertainty: Low, Middle and High. The percent of correct classification (PCC) for the test set is used as a criterion to judge the performance of our classifier. The results summarized in Table 1 show that the belief rough set classifier works very well in certain and uncertain case for all databases with the different degrees of uncertainty. However, when the degree of uncertainty increases there is a slight decline in accuracy.

 Table 1. Experimental results

	PCC	PCC	PCC	PCC
Database	certain case	Low Unc	Middle Unc	High Unc
Wisconsin breast cancer database	82.61%	81.2%	79.53%	74.93%
Balance scale database	73.7%	72.3%	70.9%	59.34%
Congressional voting records database	94.11%	93.74%	92.53%	82.53%

6 Conclusion

In this paper, we have developed a new rough set classifier to handle uncertainty in decision attributes of databases. The belief rough set classifier is shown to generate significant belief decision rules for standard classification datasets. The proposed algorithm is computationally expensive. So, our future work will develop heuristics for simplification of the decision system for efficient generation of significant decision rules without costly calculation.

References

- 1. Bosse, E., Jousseleme, A.L., Grenier, D.: A new distance between two bodies of evidence. Information Fusion 2, 91–101 (2001)
- Grzymala-Busse, J.: Rough set strategies to data with missing attribute values. In: Workshop Notes, Foundations and New Directions of Data Mining, the 3rd International Conference on Data Mining, Melbourne FL, USA, pp. 56–63 (2003)

 $^{^1~{\}rm http://www.ics.uci.edu/~mlearn/MLRepository.html}$

- 3. Pawlak, Z., Zdzisław, A.: Rough Sets: Theoretical Aspects of Reasoning About Data. Kluwer Academic Publishing, Dordrecht (1991)
- 4. Shafer, G.: A mathematical theory of evidence. Princeton University Press, Princeton (1976)
- Slowinski, R., Stefanowski, J.: Rough classification in incomplete information systems. Mathematical and Computer Modeling 12, 1347–1357 (1989)
- Smets, P., Kennes, R.: The transferable belief model. Artificial Intelligence 66, 191– 236 (1994)
- Trabelsi, S., Elouedi, Z.: Attribute selection from partially uncertain data using rough sets. In: The Third International Conference on Modeling Simulation, and Applied Optimization, Sahrjah, U.A.E., January 20-22 (2009)