PROBABILISTIC ROUGH COMPUTING UNDER FUZZINESS USING THRESHOLDS

B.Krishnaveni[1], G.Ganesan[2]

Abstract: Rough Sets, an invention of Z.Pawlak finds various applications in Knowledge Representation, Knowledge Acquisition, Data Mining etc. Considering the importance of this theory, various researchers have been paying attention in introducing fuzzy and intuitionistic fuzzy concepts. In 2012, G.Ganesan has discussed the Probabilistic rough set approaches of Yao under fuzzy environment. In this paper we extended the concepts of Naïve Bayesian Rough Sets Model under fuzziness using two thresholds

***Keywords:** fuzziness, probabilistic rough set, rough set*

1. INTRODUCTION

The theory of rough sets by Pawlak has been playing vital role in the areas such as knowledge discovery, data mining, information retrieval etc.In this model, Pawlak approximated each input or concept in terms of union of basic categories in to two ways. However this approach does not reveal the degree of involvement of the basic categories under consideration. Hence W.Ziarko introduced variable precession rough sets in 1993,which spells the degrees of contributions of the basic categories while approximating a given input. Later Bing Zhou,YY Yao,Slezak etc have been working in extending VPRS model through probabilistic approaches.

Considering the importance of rough computing Dubios and Prade have worked on hybridizing rough and fuzzy models for various applications. In 2005,G.Ganesan et.al., introduced the concept of thresholds in rough fuzzy computing. In 2012,G.Ganesan et.al., have discussed the Naive Bayesian rough set model under fuzziness [3] through a threshold ,by considering the importance of Naive Bayesian rough set model which was discussed in [8,9].

The paper consists of seven sections; Section 2 deals with the introductory concepts of Pawlak and subsequent contributions by others; Section 3 deals with describing the rough computing concepts of thresholds for fuzzy concept; Section 4 deals with designing a Naïve Bayesian Model for a fuzzy input under a threshold. In section 5 and 6, we describe the concepts of Two thresholds for a fuzzy input and the corresponding modifications in designing a Naïve Bayesian Model for a fuzzy input under two thresholds.

2. PROBABILISTIC ROUGH SETS

The Rough Sets [5,6] theory gives two way approximations namely lower and upper approximations for a given input. For given finite universe of discourse U and an equivalence relation E, we define the equivalence class of any $x \in U$ to be

$[x] = \{ y \in U / xEy \}$. The family of equivalence classes $U/E = \{ [x]_E \mid x \in U \}$ is a partition of the universe U.

For a given concept C, according to Pawlak, the lower approximation is $\underline{apr}_E(C) = \{ x \in U / [x]_E \subseteq C \}$ and upper approximation is $\overline{apr}_E(C) = \{ x \in U / [x]_E \cap C \neq \Phi \}$. According to him, for a given concept C, we obtain three disjoint regions namely positive, negative and boundary regions which are defined as follows:

Positive Region: $POS_E(C) = \{ x \in U / [x]_E \subseteq C \}$

Boundary : $BND_E(C) = \{ x \in U / [x]_E \cap C \neq \Phi \wedge [x]_E \not\subseteq C \}$

Negative region: $NEG_E(C) = \{ x \in U / [x]_E \cap C = \Phi \}$

Since Pawlak's model is restrictive and lacks with quantifying the grade of the involvement of Basic Categories, several researchers focused on generalizing this approach towards parameterized rough sets, probabilistic rough sets, generalized rough sets etc.

In 1994, Pawlak and Skowron [7] defined rough membership function by considering degrees of overlap between equivalence classes and a concept C to be approximated and is viewed as the conditional probability of an object belongs to C given that the object is in [x] (for simplicity, we denote $[x]_E$ with [x]) which is given as $\Pr\left(C/[x] \right) = \frac{|C \cap [x]|}{|[x]|}$

Using the definition quoted above, in [9], the positive, boundary and negative regions are defined as follows:

$$POS(C) = \left\{ x \in U / \Pr\left(C/[x] \right) = 1 \right\}$$

$$BND(C) = \left\{ x \in U / 0 < \Pr\left(C/[x] \right) < 1 \right\}$$

$$NEG(C) = \left\{ x \in U / \Pr\left(C/[x] \right) = 0 \right\}$$

In 2009, Greco et.al [4] introduced parameterized rough sets model by generalizing the above said definitions. In this model, two thresholds namely α and β are used to define the probabilistic regions and using them, the positive, boundary and negative regions are modified as follows:

$$POS_{(\alpha,\beta)}(C) = \left\{ x \in U \;/\; \Pr\left(C/[x] \right) \geq \alpha \right\}$$

$$BND_{(\alpha,\beta)}(C) = \left\{ x \in U \;/\; \beta < \Pr\left(C/[x] \right) < \alpha \right\}$$

$$NEG_{(\alpha,\beta)}(C) = \left\{ x \in U \;/\; \Pr\left(C/[x] \right) \leq \beta \right\}$$

These Probabilistic regions will lead three way decisions namely acceptance, deferment and rejection respectively for any object x in U. But, however, in several cases, it is easy to compute the probability of the existence of a category [x] for a given concept C using $\Pr\left([x]/C \right) = \frac{|[x] \cap C|}{|C|}$

Hence, by Baye's Theorem, in [9], the Positive, Boundary and Negative Regions are given by

$$POS^{B}_{(\alpha',\beta')}(C) = \left\{ x \in U \;/\; \log \frac{\Pr([x]/C)}{\Pr([x]/C^{c})} \geq \alpha' \right\}$$

$$BND^{B}_{(\alpha',\beta')}(C) = \left\{ x \in U \;/\; \beta' < \log \frac{\Pr([x]/C)}{\Pr([x]/C^{c})} < \alpha' \right\}$$

$$NEG^{B}_{(\alpha',\beta')}(C) = \left\{ x \in U \;/\; \log \frac{\Pr([x]/C)}{\Pr([x]/C^{c})} \leq \beta' \right\}$$

$$where \quad \alpha' = \log \frac{\Pr(C^{c})}{\Pr(C)} + \log \frac{\alpha}{1-\alpha}$$

$$and \quad \beta' = \log \frac{\Pr(C^{c})}{\Pr(C)} + \log \frac{\beta}{1-\beta}$$

Now, we shall discuss the conventional approach on dealing the fuzzy sets to approximate under rough computing, which was discussed in [1]

3. ANALYSIS OF FUZZY SET USING A THRESHOLD

Consider a set D, called **R-domain** [1], satisfying the following properties:

a) $D \subset (0,1)$

b) If a fuzzy concept C is under computation, eliminate the values $\mu_C(x)$ and $\mu_{C^c}(x)$ $\forall x \in U$ from the domain D, if they exist.

ISBN - 978-93-81583-56-2

c) After the computation using C, the values removed in (b) may be included in D provided C must not involve in further computation

Consider the universe of discourse $U=\{x_1,x_2,\ldots,x_n\}$. Let $\alpha,\alpha_1,\alpha_2,\beta$ be the thresholds assume one of the values from the domain D, where D is constructed using the fuzzy concepts A and B. For a given threshold α and a fuzzy set A, the Strong α-Cut is given by $A[\alpha]=\{x\in U / \mu_A(x)>\alpha\}$. The union and intersection of fuzzy sets [10] are by the maximum and minimum of corresponding membership values respectively.

Using the definitions of fuzzy sets mentioned above, the following properties were derived in [1].

a. $A[\alpha_1]\cup A[\alpha_2]=A[\alpha]$ where $\alpha=\min(\alpha_1,\alpha_2)$
b. $A[\alpha_1]\cap A[\alpha_2]=A[\alpha]$ where $\alpha=\max(\alpha_1,\alpha_2)$
c. $(A\cup B)[\alpha]=A[\alpha]\cup B[\alpha]$
d. $(A\cap B)[\alpha]=A[\alpha]\cap B[\alpha]$
e. $A^c[\alpha]=A[1-\alpha]^c$
f. $(A\cup B)^c[\alpha]=A^c[\alpha]\cap B^c[\alpha]$
g. $(A\cap B)^c[\alpha]=A^c[\alpha]\cup B^c[\alpha]$

Using the mathematical tool derived as above, in [1], rough set approach on fuzzy sets using a threshold is introduced as discussed below.

<u>3.1 Rough Approximations on fuzzy sets using α:</u> Let Ψ be any partition of U, say $\{B_1,B_2,\ldots,B_t\}$. For the given fuzzy concept, the lower and upper approximations with respect to α can be defined as ${}_{\alpha}C=\underline{(C[\alpha])}$ and ${}^{\alpha}C=\overline{(C[\alpha])}$ respectively.

3.1.1 Propositions

Here, by using the properties of rough sets, the following propositions [1] can be obtained.

a. ${}^{\alpha}(A\cup B)={}^{\alpha}A\cup{}^{\alpha}B$

b. ${}_{\alpha}(A\cap B)={}_{\alpha}A\cap{}_{\alpha}B$

c. ${}_{\alpha}(A\cup B)\supseteq{}_{\alpha}A\cup{}_{\alpha}B$

d. ${}^{\alpha}(A\cap B)\subseteq{}^{\alpha}A\cap{}^{\alpha}B$

e. ${}^{\alpha}(A^c)=({}_{1-\alpha}A)^c$

f. ${}_{\alpha}(A^c)=({}^{1-\alpha}A)^c$

Now, we shall hybridize the concepts dealt in the above two sections which gives the approach of dealing a fuzzy concepts under Naïve Bayesian Probabilistic Rough Sets as discussed in [3].

4. NAÏVE BAYESIAN PROBABILISTIC ROUGH SETS MODEL FOR A FUZZY CONCEPT

Since, in the above both sections, the same threshold α has been used, for different purposes, to make the homogeneity, in this paper, we replace the threshold α to obtain a Strong Cut on fuzzy sets with δ.

Hence, for a given fuzzy concept F with the threshold δ, the probabilistic positive, boundary and negative regions are respectively defined on the approximation space U/E as

$$POS_{\delta}(F) = \left\{ x \in U / \Pr\left(F[\delta] / [x] \right) = 1 \right\}$$

$$BND_{\delta}(F) = \left\{ x \in U / 0 < \Pr\left(F[\delta] / [x] \right) < 1 \right\}$$

$$NEG_{\delta}(F) = \left\{ x \in U / \Pr\left(F[\delta] / [x] \right) = 0 \right\}$$

For given parameters α and β, the regions of the parameterized rough sets model are given by

$$POS_{(\alpha,\beta,\delta)}(F) = \left\{ x \in U / \Pr\left(F[\delta] / [x] \right) \geq \alpha \right\}$$

$$BND_{(\alpha,\beta,\delta)}(F) = \left\{ x \in U / \beta < \Pr\left(F[\delta] / [x] \right) < \alpha \right\}$$

$$NEG_{(\alpha,\beta,\delta)}(F) = \left\{ x \in U / \Pr\left(F[\delta] / [x] \right) \leq \beta \right\}$$

and the Regions of Naïve Bayesian Rough Sets Model are given by

$$POS^{B}_{(\alpha',\beta,\delta)}(F) = \left\{ x \in U / \log \frac{\Pr\left([x] / F[\delta] \right)}{\Pr\left([x] / (F[\delta])^{C} \right)} \geq \alpha' \right\}$$

$$BND^{B}_{(\alpha',\beta,\delta)}(F) = \left\{ x \in U / \beta' < \log \frac{\Pr\left([x] / F[\delta] \right)}{\Pr\left([x] / (F[\delta])^{C} \right)} < \alpha' \right\}$$

ISBN - 978-93-81583-56-2

$$NEG^{B}_{(\alpha',\beta',\delta)}(F)=\left\{x\in U\,/\,\log\frac{\Pr\left(\left[x\right]/F[\delta]\right)}{\Pr\left(\left[x\right]/(F[\delta])^{C}\right)}\le\beta'\right\}$$

$$where\quad \alpha'=\log\frac{\Pr(C^{c})}{\Pr(C)}+\log\frac{\alpha}{1-\alpha}\quad and\quad \beta'=\log\frac{\Pr(C^{c})}{\Pr(C)}+\log\frac{\beta}{1-\beta}$$

5. ANALYSIS OF FUZZY SET USING TWO THRESHOLDS:

Consider the universe of discourse $U=\{x_1,x_2,\ldots,x_n\}$. Let $\alpha,\alpha_1,\alpha_2,\beta_1,\beta_2,\beta$ be the thresholds assume one of the values from the domain D where D is constructed using the fuzzy sets A and B.

Define $A[\alpha,\beta]=\{x\in U/\ \alpha<\mu_A(x)\le\beta\}$ where $\alpha<\beta$. However, by the definition of D, $A[\alpha,\beta]$ can also be defined as

$A[\alpha,\beta]=\{x\in U/\ \alpha<\mu_A(x)<\beta\}$
$A[\alpha,\beta]=\{x\in U/\ \alpha\le\mu_A(x)\le\beta\}$
$A[\alpha,\beta]=\{x\in U/\ \alpha\le\mu_A(x)<\beta\}$

The properties of these thresholds were discussed in [2] as given below:

5.1 Properties

a. $A[\alpha_1,\beta_1]\cup A[\alpha_2,\beta_2]=A[\alpha,\beta]$ where $\alpha=\min(\alpha_1,\alpha_2)$ and $\beta=\max(\beta_1,\beta_2)$

b. $A[\alpha_1,\beta_1]\cap A[\alpha_2,\beta_2]=A[\alpha,\beta]$ where $\alpha=\max(\alpha_1,\alpha_2)$ and $\beta=\min(\beta_1,\beta_2)$

c. $(A\cup B)[\alpha,\beta]\subseteq A[\alpha,\beta]\cup B[\alpha,\beta]$

d. $(A\cap B)[\alpha,\beta]\supseteq A[\alpha,\beta]\cap B[\alpha,\beta]$

e. $A^c[\alpha,\beta]=A[1-\beta,1-\alpha]$

f. $(A\cup B)^c[\alpha,\beta]\subseteq A^c[\alpha,\beta]\cap B^c[\alpha,\beta]$

g. $(A\cap B)^c[\alpha,\beta]\supseteq A^c[\alpha,\beta]\cup B^c[\alpha,\beta]$

5.2. Rough set approach on fuzzy sets using α and β: Let X be any partition of U, say $\{B_1,B_2,\ldots,B_t\}$. For the given fuzzy set A, the lower and upper approximations with respect to α and β can be defined as $A_{\alpha,\beta}=\underline{(A[\alpha,\beta])}$ and $A^{\alpha,\beta}=\overline{(A[\alpha,\beta])}$ respectively.

Here, as quoted in [2], by using the properties of rough sets, the following propositions can be obtained.

1. $(A\cup B)^{\alpha,\beta} \subseteq A^{\alpha,\beta} \cup B^{\alpha,\beta}$
2. $(A\cap B)_{\alpha,\beta} \supseteq A_{\alpha,\beta} \cap B_{\alpha,\beta}$
3. $(A^c)^{\alpha,\beta} = A^{1-\beta,1-\alpha}$
4. $(A^c)_{\alpha,\beta} = A_{1-\beta,1-\alpha}$

5.2.1. Example: Consider the universe of discourse U={a,b,c,d,e,f} with the partition X={{a},{,b}, {c}, {d,e,f}}. Let A=(0.2,0.5,0.3,0.5,0.7,0) and B=(0.6,0.8,1, 0.4, 0.6, 0.6). Then D=(0,1)-{0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8}. Let $\alpha,\beta \in D$, say, α=0.45 and β=0.75

Then $A[\alpha,\beta]$={b,d,e} and $B[\alpha,\beta]$={a,e,f}. Hence $A_{\alpha,\beta}=\phi$; $A^{\alpha,\beta}$={b,d,e,f}; $B_{\alpha,\beta}=\phi$ and $B^{\alpha,\beta}$={a,d,e,f} $A\cup B$= (0.6, 0.8, 1, 0.5, 0.7, 0.6); $(A\cup B)[\alpha,\beta]$={a,d,e,f}. Hence $(A\cup B)^{\alpha,\beta}$={a,d,e,f}. But, $A^{\alpha,\beta}\cup B^{\alpha,\beta}$={a,b,d,e,f}. Hence, $(A\cup B)^{\alpha,\beta} \subseteq A^{\alpha,\beta}\cup B^{\alpha,\beta}$

$A\cap B$=(0.2, 0.5, 0.3, 0.4, 0.6, 0); $(A\cap B)[\alpha,\beta]$={b,e}. Hence $(A\cap B)_{\alpha,\beta}$={b}. But, $A_{\alpha,\beta}\cap B_{\alpha,\beta} = \Phi$. Hence, $(A\cap B)_\alpha \supseteq A_\alpha \cap B_\alpha$.

Now, A^c=(0.8,0.5,0.7,0.5,0.3,1). $A^c[\alpha,\beta]$= {b,c,d}. Hence $(A^c)_{\alpha,\beta}$={b,c} and $(A^c)^{\alpha,\beta}$={b,c,d,e,f}.Now, $A[1-\beta,1-\alpha]$= {b,c,d}. Hence, $(A^c)_{\alpha,\beta}=A_{1-\beta,1-\alpha}$ and $(A^c)^{\alpha,\beta}= A^{1-\beta,1-\alpha}$.

Now, we introduce Naïve Bayesian Probabilistic Rough Sets Model on Fuzzy sets using two thresholds

6. NAÏVE BAYESIAN PROBABILISTIC ROUGH SETS MODEL FOR A FUZZY CONCEPT WITH TWO THRESHOLDS

Similar to one threshold approach, here also we replace the threshold α and β with δ and π. Hence, for a given fuzzy concept F with the thresholds δ and π, the probabilistic positive, boundary and negative regions are respectively defined on the approximation space U/E as

$$POS_{(\delta,\pi)}(F) = \left\{ x \in U \,/\, \Pr\left(F[\delta,\pi] \big/ [x] \right) = 1 \right\}$$

$$BND_{(\delta,\pi)}(F) = \left\{ x \in U \,/\, 0 < \Pr\left(F[\delta,\pi] \big/ [x] \right) < 1 \right\}$$

$$NEG_{(\delta,\pi)}(F) = \left\{ x \in U \,/\, \Pr\left(F[\delta,\pi] \big/ [x] \right) = 0 \right\}$$

For given parameters α and β, the regions of the parameterized rough sets model are given by

ISBN - 978-93-81583-56-2

$$POS_{(\alpha,\beta,\delta,\pi)}(F)=\left\{x\in U \,/\, \Pr\left(F[\delta,\pi]\big/[x]\right)\geq \alpha\right\}$$

$$BND_{(\alpha,\beta,\delta,\pi)}(F)=\left\{x\in U \,/\, \beta<\Pr\left(F[\delta,\pi]\big/[x]\right)< \alpha\right\}$$

$$NEG_{(\alpha,\beta,\delta,\pi)}(F)=\left\{x\in U \,/\, \Pr\left(F[\delta,\pi]\big/[x]\right)\leq \beta\right\}$$

and the Regions of Naïve Bayesian Rough Sets Model are given by

$$POS^{B}_{(\alpha',\beta',\delta,\pi)}(F)=\left\{x\in U \,/\, \log\frac{\Pr\left([x]/F[\delta,\pi]\right)}{\Pr\left([x]/(F[\delta,\pi])^{C}\right)}\geq \alpha'\right\}$$

$$BND^{B}_{(\alpha',\beta',\delta,\pi)}(F)=\left\{x\in U \,/\, \beta'<\log\frac{\Pr\left([x]/F[\delta,\pi]\right)}{\Pr\left([x]/(F[\delta,\pi])^{C}\right)}< \alpha'\right\}$$

$$NEG^{B}_{(\alpha',\beta',\delta,\pi)}(F)=\left\{x\in U \,/\, \log\frac{\Pr\left([x]/F[\delta,\pi]\right)}{\Pr\left([x]/(F[\delta,\pi])^{C}\right)}\leq \beta'\right\}$$

$$\textit{where}\quad \alpha'=\log\frac{\Pr(C^{c})}{\Pr(C)}+\log\frac{\alpha}{1-\alpha}\quad \textit{and}\quad \beta'=\log\frac{\Pr(C^{c})}{\Pr(C)}+\log\frac{\beta}{1-\beta}$$

7. CONCLUSION

In this paper, we introduced the concepts of thresholds in modeling Naïve Bayesian Rough Sets which helps the researchers in making various applications.

8. REFERENCES

1. G.Ganesan et. al., Rough Set: Analysis of Fuzzy Sets using thresholds, UGC Sponsored National Conference on Recent Tends in Computational Mathematics, March 2004, Gandhigram Rural Institute, Tamilnadu, by Narosa Publishers, pp:81-87, 2005
2. G.Ganesan et. al., Rough index induced by a fuzzy set using two thresholds, International Conference on Soft Computing, pp:164-169,BIHER, Chennai, May 2004
3. G.Ganesan et. al., Naïve Bayesian Rough Sets under Fuzziness, International Journal of Computer, Mathematical Sciences & Applications [To appear]
4. Greco, S., Matarazzo, B. and Slowinski, R. Parameterized rough set model using rough membership and Bayesian con_rmation measures, International Journal of Approximate Reasoning, 49, 285-300, 2009

5. Pawlak, Z. Rough sets, International Journal of Computer and Information Sciences, 11, 341-356, 1982.
6. Pawlak, Z. Rough Sets, Theoretical Aspects of Reasoning about Data, Dordrecht: Kluwer Academic Publishers, 1991.
7. Pawlak, Z. and Skowron, A. Rough membership functions, in: Yager, R.R., Fedrizzi, M. and Kacprzyk, J., Eds., Advances in the Dempster-Shafer Theory of Evidence,John Wiley and Sons, New York, 251-271, 1994.
8. Slezak, D. and Ziarko, W. The investigation of the Bayesian rough set model, International Journal of Approximate Reasoning, 40, 81-91, 2005.
9. Yiyu Yao and Bing Zhou, Naive Bayesian Rough Sets, Proceedings of RSKT 2010, LNAI 6401, pp. 719-726, 2010.
10. Zadeh L.A. Fuzzy Sets, Journal of Information and Control, 8, pp. 338-353, 1965

* * * *

----

[1]B.Krishnaveni, Department of Mathematics, Adikavi Nannaya University, Rajahmundry, Andrha Pradesh, India Email: krishnavenibadam@yahoo.co.in
[2]G.Ganesan, Department of Mathematics, Adikavi Nannaya University, Rajahmundry, Andrha Pradesh, India Email: prof.ganesan@yahoo.com