

# INFORMATION FUSION BASED ON MODIFIED PROPORTIONAL CONFLICT REDISTRIBUTION RULES FOR CARDIO VASCULAR DISEASE PREDICTION

T. Marikani<sup>1\*</sup>, K. Shyamala<sup>2</sup>

<sup>1</sup>  
Assistant Professor, Dept. of Computer Science,  
Sree Muthukumaraswamy College, Affiliated to University of Madras,  
Chennai, Tamilnadu, India  
tmkani330@gmail.com

<sup>2</sup>  
Associate Professor, P.G & Research Dept. of Computer Science,  
Dr. Ambedkar Govt.Arts College(Autonomous), Chennai, Tamilnadu,  
India  
shyamalakannan2000@gmail.com

## *Abstract*

Cardiovascular sickness is a foremost rationale of bleakness and mortality in the current livelihood. Healthcare industry is commonly "rich in information", but regrettably all the data are not mined [4]. Data mining is most essential for discovering the hidden patterns and to take the effective decision based upon their algorithms and techniques. In this paper we proposed Modified Proportional Conflict Redistributions fusion rules for the combination of uncertainty and conflicting information, which allow us to deal with highly conflicting data which is available in the database.

Modified Proportional Conflict Redistribution1 (PCR1) rule is one of the efficient fusion rules for combination of indecisive, vague and highly inconsistent data. Conjunctive rule has been applied in the proposed work. The proposed Modified PCR1 mechanism is applied in the Cleveland dataset which is obtained from UCI machine learning repository and obtained 88.5% accuracy. The experimental consequences show substantiation of the performance of the proposed strategy in prediction of the cardio vascular diseases is high in terms of accuracy which makes the Modified PCR1 method is suitable for coeval applications too.

## **Keyword**

Data Mining, Modified Proportional Conflict Redistribution1 (PCR1), Classification Algorithm, Conjunctive Rule, Fusion Rule.

## **1. Introduction**

Cardiovascular diseases are one of the large amount lethal ailment across the sphere. Clinically, to check the therapeutic stipulation of a human heart, stethoscope is used. Heart condition of the patients can be understand and interpret only by the trained medical professionals.

Basically varieties of malady affect a diacritic. Some of the affliction is called congenital diseases, which occur for a solitary, right from their birth. Some the indisposition occurs owing to their heredity. Cardiovascular disease is one among the disease, which has the possibilities of phenomenon amid the people due to their heredity. Nowadays modern hospitals are well outfitted with monitoring and other data

collection devices ensuing in massive data which are collected incessantly through health assessment and medical treatment.

This paper presents a machine learning based classification system, to predict the heart disease among the patients. This can be predicted based on few attributes like age, chest pain type, cholesterol, fasting blood sugar, resting electrocardiographic results, exercise induced angina and resting blood pressure [11]. The proposed framework involves prediction of occurrence of heart disease among the patients based on proportional conflict redistribution (PCR1) rule. The PCR1 is the preponderance efficient and effectual redistribution rule, which has been successfully implemented to many problems.

## 2. Literature Work

Florentin Smarandache et.al. [1], proposed the five version of Proportional Conflict Re- distribution rule (PCR) for information fusion together with numerous examples. In PCR1 the author restricted from the hyper-power set to the power set and without degenerate cases. The PCR rule redistribute the conflicting mass, after the conjunctive rule has been applied, proportionally with some functions depending on the masses assigned to their corresponding columns in the mass matrix. DSMT is an extension of Bayesian theories. DSMT theories have proficient combination of vague, indefinite and highly conflicting sources of information. In this work, the author describe from PCR1 to PCR5. PCR1 basically compute the total conflicting mass. The total mass which is involved in the conflict is distributed to the non-empty set in PCR2. In PCR3, partial conflict involved. That is, instead of using the total conflicting mass, one transfers partial conflicting masses to non-empty sets. PCR4 redistributes the partial conflicting mass to the elements implicated in the partial conflict. PCR5 increases the intricacy of the rules and also the precision of the redistribution of conflicting masses.

F. Smarandache et.al. [2], performed work on PCR1 along with WAO (Weighted Average Operator). Dynamic fusion is introduced by author in this proposed work. The feature of PCR1 is that, it works on degenerate cases as well as on non-degenerate cases. Author applied the proposed work on Smets' rule, Yager's rule, DSm hybrid rule, Murphy's rule, Dubois-Prade's rule and Dempster's rule. Basically the entire rule uses the combination of commutative and associative rule. The hybrid DSm Rule is a new powerful rule of combination. The author widen WAO for the degenerate case, the conflict mass is transferred to non-empty disjunctive form when all non-empty sets together, i.e., when all column sums of all non empty sets are zero.

Purushottam et.al. [3], proposed the exertion on efficient heart disease prediction system using data mining techniques. The author used test and training data to find out the accuracy of the dataset. K-10 fold method is used to test and training dataset and find the accuracy of 86.3% and 87.3% respectively. Compare to other mining techniques this method provide the better results, which help medical practitioner in taking effective decision making about the heart disease level among the patients. Some of the rules generated by proposed systems are original rule, pruned rule and classified rule, rules without duplicates, polish and sorted rule.

Sentz K [4], author explained study of Dempster-Shafer theory offers an option to customary probabilistic theory for the mathematical representation of ambiguity. Author analyzed the three imperative functions in Dempster-Shafer theory. At the foremost stage, analyzed the basic probability assignment function called as bpa. In the second stage, the Belief function (Bel) is analyzed and at the last stage the Plausibility function (Pl) is investigated. Different types of rules are examined in this paper like Yager's Rule, Inagaki's Rule, Zhang's Rule and so on. This theory effectively combines the Bayesian notion of probabilities with the traditional idea of sets. Author explains Dempster-Shafer theory does not dictate a proclamation regarding the probability of the individual constituents of the set or interval.

## 3. Methodologies Used

### 3.1 Fusion

Fusion is assimilation of information that formed from numerous sources and demoralized that conjoining or merging facts from various tasks such as answering questions and decision making. Fusions are basically two types. Static fusion and dynamic fusion.

In static fusion all the belief functions are combined simultaneously.

In dynamic fusion the belief functions are available sequential one subsequent to another, and the current belief function is restructured by combining itself with a recently obtainable belief function.

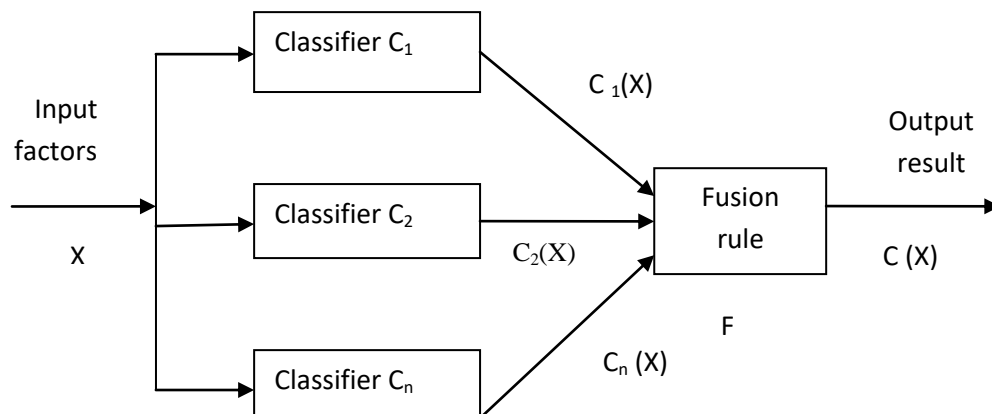


Figure 1. Fusion theory

The above figure evidently reveals that, we can have input factors and 'n' number of classifiers like  $C_1$ ,  $C_2$ ,  $C_3, \dots, C_n$  in a fusion. All the input factors are combined together and undergo various classifiers stages and finally the classified data's are analyzes in the fusion rule and produce the result.

### 3.2 Dezert-Smarandache Theory-DSMT

The Dezert-Smarandache Theory (DSmT) is well thought-out as an expansion of the Dempster-Shafer (DS) on top of the Bayesian theories. Since, Bayesian theories deal with statistical classifier, whereas DSmT deals with fusion rule. DSmT is an substitute of mathematical probabilistic theory. The fusion process for proficient grouping of uncertain, imprecise and highly conflicting sources of consel. Dempster's rule transfers the total conflicting mass to non-empty sets proportionally with their resulting stacks [2].

In this paper we modified Proportional Conflict Re-distribution rule (PCR) for information fusion together with heart disease dataset which is available from UCI Cleveland dataset. PCR1 constrained from the hyper-power set to the power set and devoid of degenerate cases gives the same consequence as the Weighted Average Operator (WAO) [3]. The PCR rules redistribute the conflicting mass, after the conjunctive rule has been applied, proportionally with some functions depending on the masses assigned to their corresponding columns in the heap matrix. The subsequent figure 2. Clearly explains the application of DSmT at different level.

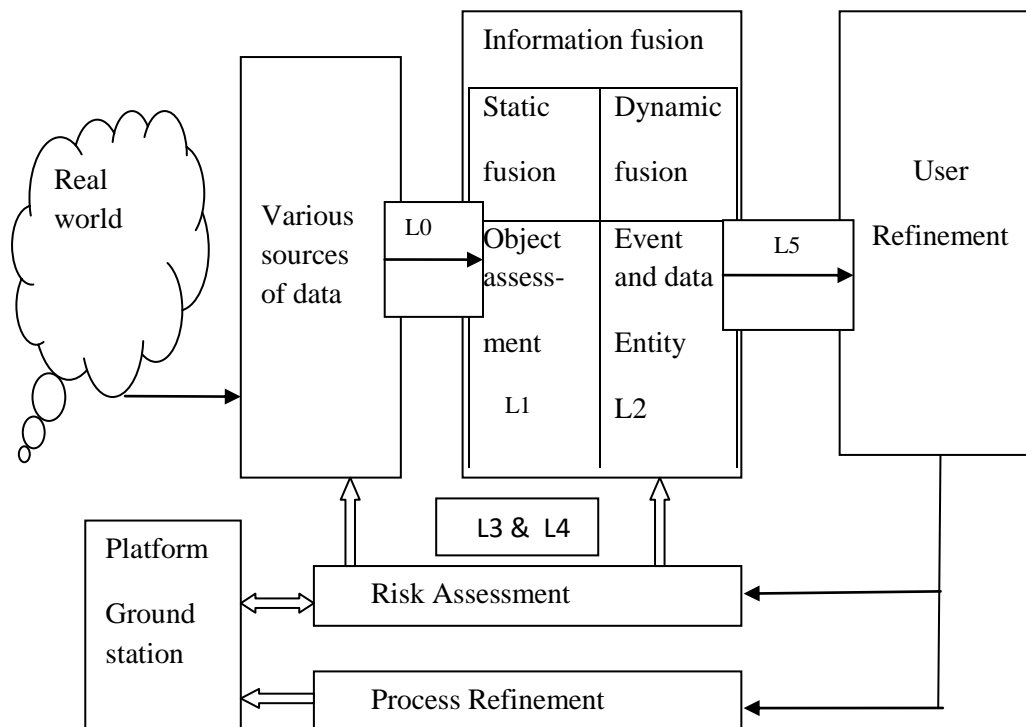


Figure 2. DSMT for Information fusion

From the above diagram, it is palpably stated that DSMT application is categorized as different level. The levels are ranged from level 0 to level 5. In level 0, data assessment is carried out i.e., the present condition of the collected data and analysis is carried out. In level 1, object assessment is prepared, that is simultaneous track and identification of data as well as image fusion is done. In level 2, situation assessment is performed i.e., event, association among data and entity determination is performed. In level 3 and 4, impact assessment like risk assessment and process refinement like performance evaluation and management are made simultaneously. In the last stage user refinement is gone through i.e., decision making process is carried out.

### 3.3 Proportional Conflict Redistribution Rule (PCR rule)

Proportional Conflict Redistributions (PCR) permits us to covenant with extremely contradictory sources for static and dynamic fusion applications [1]. It deals with ambiguity and incompatible information which is available in the data. Conjunctive rule has been applied on proportional conflict redistributions. The common theory of PCR rules is to apply the conjunctive rule on the dataset and then redistribute the conflicting masses assigned to their equivalent columns in the mass matrix.

PCR rules are hinge on total or partial conflicting masses, which are moved to the consequent sets proportionally to some functions based on weighting coefficients, which depend on their corresponding mass matrix columns [3]. It relocates conflicting masses to non-empty sets implicated in the conflicts proportionally with reverence to the masses assigned to them. At the beginning stage, PCR calculate the conjunctive rule of the belief masses of sources. Next, it calculate the composite or sectional conflicting masses and redistribute it to the non-empty set, which is involved in the conflicts masses.

#### Advantage of PCR1

- The foremost advantage is that, it can calculate both the total or partial conflict masses.

- The execution of PCR1 rule is very easy and simple.
- It is less cost effective and an easy alternative fusion rule for an existing rule.
- Very easy to adoptable to PCR1, as it contain simple formula and algorithm.
- As the specificity of knowledge is preserved, PCR1 rule mechanism works fine compare to other rules.
- Reflects the majority rules compare to other rules.

**3.4 Conjunctive Rule**

If both sources of information are telling the truth, then we apply the conjunctive rule [4], which means consensus between them:

$$\forall A \in S^\theta, \text{ one has } m_1(A) = \sum_{\substack{X_1, X_2 \in S^\theta \\ X_1 \cap X_2 = A}} m_1(X_1) m_2(X_2)$$

Where the Total Conflicting Mass is:

$$K_{12} = \sum_{\substack{X_1, X_2 \in S^\theta \\ X_1 \cap X_2 = \emptyset}} m_1(X_1) m_2(X_2)$$

**3.5 Qualitative Conjunctive Rule**

The Qualitative Conjunctive Rule (qCR) of  $s \geq 2$  sources is defined similarly to the quantitative conjunctive consensus rule, i.e.

$$q m_{qCR}(X) = \sum_{\substack{X_1, \dots, X_n \in G^\theta \\ X_1 \cap \dots \cap X_n = X}} \prod_{i=1}^s q m_i(X_i)$$

The total qualitative conflicting mass is given by

$$K_{1, \dots, s} = \sum_{\substack{X_1, \dots, X_n \in G^\theta \\ X_1 \cap \dots \cap X_n = \emptyset}} \prod_{i=1}^s q m_i(X_i)$$

**3.6. General Principles of PCR rule**

Let's  $\theta = \{ \theta_1, \theta_2, \dots, \theta_n \}$  be the frame of the fusion problem under consideration and two belief assignments  $m_1, m_2 : G \rightarrow [0,1]$  such that  $\sum_{x \in G} m_i(x) = 1$ , where  $i=1,2$ . The following are the some of the steps to be proceeded for PCR as follows:

- Step 1: Compute the conjunctive rule,  $\forall X \in G$ .
- Step 2: Compute the partial or total conflicting masses.
- Step 3: Proportionally redistribute the conflicting partial or total mass to non-empty sets.

**4. Algorithm for Modified PCR1**

**Algorithm 1: Modified Proportional Conflict Redistribution1****Input : Proposition, positive and negative proposition, positive and negative mass, mass\_pcr****Output : Accurate predict percentage**


---

```

1. Begin
2.   Algorithm pcr(data)
3.   proposition=findpropositions(data)
4.     positive_proposition=calculate_positive_proposition(data)
5.     conflict_proposition=calculate_conflict_proposition(data)
6.     positive_mass=calculate_positive_mass(data,positive_proposition)
7.     conflict_mass=calculate_conflict_mass(data,conflict_proposition)
8.     mass_pcr=update_mass_calculated(data,conflict_mass,positive_mass)
9.     if(mass_pcr>=6.25) :
10.       predict="positive"
11.     else:
12.       predict="negative"
13.     accuracy = calculate_accuracy(predict,dataoutcome)
14.   Return accuracy percentage
15. End

```

The above proposed Modified Proportional Conflict Redistribution algorithm helps to find out the accurate percentage for heart disease prediction among the patients, where the dataset was collected from UCI machine learning repository [11]. In the foremost stage, proposition of the data was designed. In the second stage, positive proposition and conflict proposition of the data was deliberated. Next, the positive mass and conflict mass was calculated based on the positive and conflict proposition. In the line number 8, mass value of the pcr is intended based on the data, positive mass and conflict mass. From line number 9 to 12, it check for if condition, which tries to predict the positive occurrences or negative occurrences of heart disease among the patients in the dataset. Finally, based on the data outcome from the previous condition checking, it calculates the accuracy percentage of heart disease among the patients.

## 5. Result and Discussion

There are many classification algorithms are available in data mining, such as the Classification tree, Random forest, Decision tree, Naïve Bayes algorithm, Bayesian belief network, Nearest Neighbor algorithm and so on[12]. Among them, we choose the Naïve Bayes algorithm. Naïve Bayes algorithm is one of the statistical classifier. It helps to find out the probability among the data. It is one of the simple and reliable algorithms even for large number of dataset.

Dezert-Smarandache Theory (DSmT) was the extension of Bayesian theory which deals with the fusion process for proficient combination of uncertain, imprecise and highly conflicting sources of information [5]. DSmT deals with fusion rule. As an alternative of applying a direct transfer of partial conflicts onto partial uncertainties as with DSmT, the idea behind the Proportional Conflict Redistribution (PCR) rule is to transfer (total or partial) conflicting masses to nonempty sets involved in the conflicts proportionally with respect to the masses assigned to them by sources.

Since DSmT is an extension of Bayesian theory, the proposed work PCR1 was compared with Naïve Bayes algorithm. As, our previous effort was modification of Modified Multinomial Naïve Bayes algorithm, it produce the accuracy percentage for heart disease prediction as 79.5% for the current dataset and our improved exertion i.e. PCR1 produce 88.5% of accuracy for heart disease prediction. Some other classification algorithms are correlate with the proposed work. The following table and figure help us to understand the work in clear manner.

**Table 1. Comparison of various classification algorithms**

Classification Tree	71.30
Naïve Bayes	79.50

SVM	75.90
Random Forest	73.90
Modified Multinomial Naïve Bayes	74.87
Modified PCR1	88.50

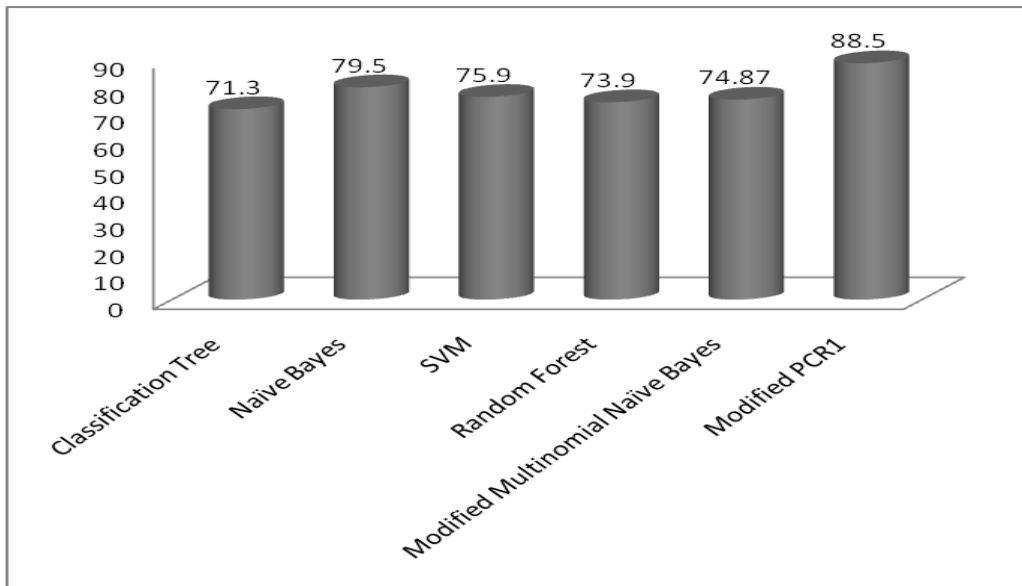


Figure 3. Comparison of various classification algorithms

## 5. Conclusion

In this paper we proposed the work on Proportional Conflict Redistribution rule (PCR1) in combination of information fusion and it is implemented by applying the conjunctive rule. Basically, PCR1 restricted from the hyper-power set to the power set without degenerating the cases. PCR1 redistribute the total conflicting mass [7]. The experimental outcome signifies that the algorithm can achieve better classification performance compared to other classification algorithm. The major contribution of this work is to help non-specialized doctors to make correct decision about the heart disease risk level prediction. The experimental consequences show substantiation of the recital of the proposed method in prediction of the cardio vascular diseases is high in terms of accuracy which makes the Modified PCR1 method is suitable for real-time applications too.

## References

1. Florentin Smarandache and Jean Dezert, "Proportional Conflict Redistribution Rules for information Fusion", Advances and Applications of DSmT for Information Fusion, ISBN: 1-59973-000-6, pp. 3-68.
2. F. Smarandache and J. Dezert, "A Simple Proportional Conflict Redistribution Rule", International Journal of Applied Mathematics & Statistics (IJAMAS) ISSN: 0973-1377, Vol. 3 (J05), pp.1-36, 2005.
3. Purushottama.C, Kanak Saxenab and Richa Sharmac, "Efficient Heart Disease Prediction System", Procedia Computer Science, ISSN No.1877-0509, pp. 962 – 969, 2016.
4. Sentz K and Ferson S., "Combination of evidence in Dempster-Shafer Theory", SANDIA Tech. Report, ISSN: 2002- 0835, 96 pages, April 2002.
5. Yager R.R., "On the Dempster-Shafer framework and new combination rules", Information Sciences, Vol. 41, pp. 93–138, 1987.
6. Smets Ph and Kennes R., "The transferable belief model", Artificial Intelligence 66 (2), pp. 191-234, 1994.

7. Zadeh L., "A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination", Artificial Intelligence Magazine 7, No.2, pp. 85-90, 1986.
8. R. Ilin and R. Deming, "Simultaneous detection and tracking of multiple objects in noisy and cluttered environment using maximum likelihood estimation framework", in OCEANS 2010 IEEE, Sydney, 2010.
9. R. Ilin, "Unsupervised learning of categorical data with competing model", Neural Networks and Learning Systems, IEEE, Vol. 23(11), pp. 1726-1737, 2012.
10. P. Sudha & A. Murugan "Assessment of proportional conflict redistribution rules for forest fire detection", International Journal of Latest Trends in Engineering and Technology, Vol 9(4), pp.115-121, 2019.
11. Uma N Dulhare, "Prediction system for heart disease using Naive Bayes and particle swarm optimization", Biomedical Research, Vol 29(12), pp.2646-2649, 2018.
12. Santhi.P and Dr.V.Murali Bhaskaran, "Performance of Classification Algorithms in Heart Disease Data", International Journal of Advanced Research and computer science, ISSN No. 0976 -5697, Vol.2 (3), pp.63-70, June 2011.
13. Daniel M., "Associativity in Combination of belief functions; a derivation of minC combination", Soft Computing, Vol 7(5), pp. 288–296, 2003.
14. Lef evre E., Colot O and Vannoorenberghe P., "Belief functions combination and conflict management", Information Fusion Journal, Elsevier Publisher, Vol. 3(2), pp. 149–162, 2002.
15. Smarandache F. and Dezert J. (Editors), "Applications and Advances of DS<sub>m</sub>T for Information Fusion", Collected Works, American Research Press, Rehoboth, June 2004, <http://www.gallup.unm.edu/smarandache/DSmT-book1.pdf>.
16. Jøsang A., Daniel M and Vannoorenberghe P., "Strategies for Combining Conflicting Dogmatic Beliefs", Proceedings of the 6th International Conference on Information Fusion, Cairns, Australia, pp. 1133–1140, 2003.
17. Smarandache F and Dezert J., "Information Fusion Based on New Proportional Conflict Re-distribution Rules", Proceedings of Fusion 2005, Philadelphia, July 26-29, 2005.
18. Smarandache F., "An In-Depth Look at Information Fusion Rules and Unification of Fusion Theories", invited speaker at NASA Langley Research Center, Hampton, VA, USA, International Journal of Applied Mathematics & Statistics, Roorkee, Vol. 2, pp.271-284, 2004.
19. Smarandache F., "An Algorithm for the Unification of Fusion Theories (UFT)", International Journal of Applied Mathematics, Vol. 20(3), pp.293-314, 2007.
20. Fixsen D and Mahler R. P. S., "The modified Dempster-Shafer approach to classification", IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans, Vol. 27(1), pp. 96–104, 1997.