



An Intelligent Hybrid Quick Reduct Particle Swarm Optimization Algorithm for Feature Reduction in Cardiac Disease Prediction

Dr. Durairaj.M¹, Sivagowry.S²

Department of Computer Science, Engineering and Technology,
Bharathidasan University, Tiruchirappalli, Tamil Nadu, INDIA.

Abstract: Medical Data Mining is an emerging interest in the field of research. Medical Data is enormous and mining the required knowledge from the data is quite complex and interesting task. Data Mining is an intellectual tool, which is used to mine the required knowledge even from imprecise, uncertainty and noisy medical data. Sometimes reducing the unnecessary features will improve the accuracy of data. Many Feature Reduction Techniques are available. Heart disease is a group of condition affecting the structure and functions of the heart and has many root causes. Heart disease is the leading cause of death in all over the world in recent years. Researchers have developed many data mining techniques for diagnosing heart disease. This paper proposes an Intelligent Hybrid Quick Reduct Particle Swarm Optimization (IHQRPSO) Algorithm which hybrids the features of Particle Swarm Optimization and Quick Reduct Algorithm for Feature Reduction in Cardiac Disease Prediction. The proposed algorithm provided promising results which supports the algorithm in Feature Reduction.

Keywords: Feature Selection, Particle Swarm Optimization, Rough Set Theory, Quick Reduct Algorithm, Hybrid, Confusion Matrix, Sensitivity, Accuracy.

I. Introduction

Feature selection is a method of eliminating features with little or no information. A large number of features with high dimension space will be large and unclean and degrade the classification accuracy [5]. In machine learning and statistic, the feature selection is considered as a problem of global combinatorial optimization [6]. Feature Selection techniques has three main advantages when constructing predictive models. The main benefits of the feature selection are that it improves the model interoperability, abridge the training time and over fitting which is for enhancing the generalization. A good feature selection mechanism is needed to amplify the processing rate with predictive exactitude and avoid incomprehensibility [7]. The problem size and the resultant search space for any learning algorithm can be reduced by using the Feature Selection. In general, two important goals of Feature Selection are to minimize the classification error and reduce the number of attributes [8]. There are number of Feature Selection algorithms available for Data Mining. The complicated task is to find the suitable algorithm which would select the feature relevant to the task without affecting the quality of the result. The Swarm Intelligent techniques and Rough Set Theory are the techniques that are mostly used for Feature reduction. But PSO is found to be effectual since it computes the individual best of each particle and finally compares all the individual particles to get the global best. It is also palpable that the Redunt feature obtained by PSO would not affect the accuracy. Rough Set also provides certain advantage in Feature reducing which does not affect the accuracy of the data set. Changes in the lifestyle, work culture and food habits lead to Cardio Vascular Disease (CVD) in developing countries. World Health Organization (WHO) reported that 30% of total global death is due to CVD [21, 22]. This estimation was done in 2008. It is also analyzed that the number of CVD Patients will increase from 300 to 600 million by 2020 [3]. By 2030, almost 25 million people will die from CVDs, mainly from heart disease and stroke [23]. Medical diagnosis is a process of decision making, during which the physician performs diagnosis of the new and unknown case from a set of clinical data from clinical experiences.

This paper is organized as follows: Section II describes about Particle Swarm optimization algorithm with its pseudo-code. Section III describes about the Rough Set Theory and Section IV gives details about the literature that support PSO and RST in medical data. Section V describes the data set used. Section VI explains the proposed algorithms followed by Results and discussion in Section VII. Section VIII concludes the paper .

II. Particle Swarm Optimization

Particle Swarm Optimization (PSO) is based on the social behaviour associated with bird's flocking for optimization problem. A social behaviour pattern of organisms that live and interact within large groups is the inspiration for PSO [20, 21, 23]. The PSO is easier to lay into operation than Genetic Algorithm. It is for the motivation that PSO doesn't have mutation or crossover operators and movement of particles is effected by

using velocity function [9]. In PSO, each particle adjusts its own flying memory and its companion's flying involvement in order to flying in the search space with velocity.

The best-fit particle of the entire swarm [10] influences the position of each particle. Each individual particle $i \in [1 \dots n]$ where $n > 1$, has current position in search space x_i , a current velocity v_i and a personal best position $P_{best,i}$ where i is the smallest value determined by objective function f . By using the $P_{best,i}$, the global best position G_{best} is calculated, which is the buck value obtained by comparing all the $P_{best,i}$.

The $P_{best,i}$ is calculated by using the formula

$$p_{best,i} = \begin{cases} P_{best,i} & \text{if } f(x_i) > P_{best,i} \\ x_i & \text{if } f(x_i) \leq P_{best,i} \end{cases} \quad \text{----} > (1)$$

The formula used to calculate Global Best Position G_{best} is

$$G_{best} = \{ \min \{ P_{best,i} \}, \text{ where } i \in [1, \dots, n] \text{ where } n > 1 \quad \text{----} > (2)$$

Velocity can be updated by using the formula

$$V_i^{t+1} = wv_i(t) + c_1r_1[x_i(t) - x_i(t) + c_2r_2[g(t) - x_i(t)] \quad \text{--} > (3)$$

where $v_i(t)$ is the velocity and w , c_1 and c_2 are user supplied co-efficient. The r_1 and r_2 are random values $x_i(t)$ is the individual best solution, $g(t)$ is the swarm's global best candidate solution. $wv_i(t)$ is known as inertia component. Inertia component value lies between 0.8 and 1.2. Lower the values of inertia component, it speeds up the convergence of swarm to optima. But higher value encourages the exploration of entire search space. $c_1r_1 [(x_i(t) - x_i(t))]$ is known as cognitive component.

III. Rough Set Theory

The concept of Rough Set was first introduced by Z.I. Pawlak in 1982 [10]. Any external parameter is not required by Rough Set Theory to analyze and conclude about data set. It is used to find the indispensable features. The discernibility matrix is used to recognize dispensable and indispensable Features [11, 24].

A. Basics

Let $I=(U,A)$ be an Information System. Here, U is an non-empty set of objects called universe and A non-empty finite set of attributes such that $a:U \rightarrow V_a$ for every $a \in A$. V_a is a set of values that attribute a may take. If $P \subseteq A$, there is an associated equivalence relation:

$$IND(P) = \{(x,y) \in U \times U \mid \forall a \in P, f_a(x) = f_a(y)\} \quad \text{-----} > 4$$

The partition of U generated by $IND(P)$ is denoted by U/P . If $(x,y) \in IND(P)$, then x and y are indiscernible by attributes from P . The equivalence classes of the P -indiscernibility relation are denoted by $[x]_P$. Let $X \subseteq U$, the P -lower Approximation $\underline{P}X$ and P -upper approximation $\overline{P}X$ of set X can be defined as:

$$\underline{P}X = \{x \mid U/[x]_P \subseteq X\} \quad \text{-----} > 5$$

$$\overline{P}X = \{x \mid U/[x]_P \cap X \neq \emptyset\} \quad \text{-----} > 6$$

Let $P, Q \subseteq A$ be equivalence relations over U , then the positive, negative and boundary regions can be defined as

$$: POS_P(Q) = \bigcup_{x \in U/Q} \underline{P}X \quad \text{-----} > 7$$

$$NEG_P(Q) = U - \bigcup_{x \in U/Q} \overline{P}X \quad \text{-----} > 8$$

$$\text{BND}_P(Q) = \bigcup_{x \in U|Q} \overline{P} X - \bigcup_{x \in U|Q} \underline{P} X \quad \xrightarrow{9} \quad \text{ides the same set of all reducts is defined as:}$$

$$\text{Red}(C) = \{ R \subseteq C \mid \gamma_R(D) = \gamma_C(D), \forall B \subseteq R, \gamma_B(D) \neq \gamma_C(D) \} \quad \xrightarrow{10}$$

A dataset may have many attribute reducts. The set of all optimal reducts is:

$$\text{Red}(C)_{\min} = \{ R \in \text{Red} \mid \forall R' \in \text{Red}, |R| \leq |R'| \} \quad \xrightarrow{11}$$

B. Quick Reduct Algorithm:

It calculates the dependency of each attribute, and the best candidate is chosen based on the calculation. The best features are added to the dataset until the dependency of the reduct candidate is equal to the consistency of dataset [13]. The below is the pseudo code of the QR Algorithm.

Algorithm QR:

```

QR(C, D)
C, the set of all conditional features;
D, the set of decision features.

R ← {}
do
    T ← R
    ∀x ∈ (C-R)
    if γR∪{x}(D) > γT(D)
        T ← R ∪ {x}
    R ← T
until γR(D) = γC(D)
return R
    
```

IV. Related Work

This section consolidates some of the related works on PSO and RST Application for extracting feature from the data set using feature selection technique.

A modified Multi Swarm PSO [8] was applied to crack discrete problems with number of sub swarm and a multi swarm scheduler that can scrutinize and control each sub swarm using the rule. For Feature Selection, Improved Feature Selection (IFS) was incorporated with Multi Swarm PSO and Support Vector Machine was integrated with F-Score. Recital of the anticipated method was compared with Standard PSO, Genetic Algorithm and Grid search. The proposed method performs drastically in terms of classification accuracy rate. In [9], a feature selection strategy is proposed which was based on rough sets and PSO. PSORSFS algorithm and other feature selection algorithms were implemented using Matlab. PSO was found to obtain optimal solution so quickly. It was observed that inertia weight and maximum velocity have an imperative brunt in the performance of PSO.

PSO is combined with K-Nearest Neighbor (PSOK-NN) classifier [14] to determine Coronary artery disease by using exercise stress testing data. It was composed of two steps. At the first step, a particle was generated which demonstrates the whole sample optimally in training data for both health and unhealthy patients. Then the class of test sample was determined by using the distance of the test sample generated by utilizing K-Nearest Neighbor algorithm. The accuracy obtained while using PSOK-NN was 92.49%.

BenxianYue and et. al., [15] explained that PSO can be used as a solution to find the appropriate features by using Rough Set Theory. The proposed method observed the change of positive region as the particles proceed through the search space. The performance of the algorithm is evaluated with Genetic Algorithm. The fitness of each individual is evaluated during search procedure. When compared with Genetic Algorithm, it is perceived that PSO requires shorter time to obtain better results.

In [8], a hybrid filter-wrapper feature subset selection algorithm based on PSO is presented. Maximum relevance minimum redundancy PSO (mr²PSO) uses the mutual information available from the filter model to weigh the bit selection probabilities in discrete PSO. This method integrated mutual information filter model

within the PSO based wrapper model. Mutual dependence of two variables is measured by the quantity called mutual information. The mutual information is expressed as a composite measure of feature relevance and redundancy. The probability of the features is weighed by using feature relevance and redundancy, which enhances the convergence rate and solution quality of feature subset selection. A stratified 10-fold cross validation is recommended.

In [19], it is recommended that hill climbing rough set approach is inadequate to find optimum solution. It is not feasible for complete search even in medium-sized data set. The PSO interacts with individuals in the population to find the optimal region in complex search space. The main pro of PSO over GA is that it doesn't require cross over or mutation operation and also inexpensive in terms of both memory and run time. Experimental results show that PSO is efficient for rough set based feature selection.

In [13], a computer aided diagnosis system has been introduced for the heart valve disease by using binary PSO and Support Vector Machine (SVM) algorithm in conjunction with K-nearest neighbor and leave-one-out cross validation. The most weighted feature is selected by using binary PSO algorithm. SVM is used to classify the outcome into two classes as healthy and having heart valve disease. The proposed method helps to optimize the feature selection process. It can also be used an ideal preprocessing tool since it increase the classification accuracy.

A new feature selection based on Rough set Theory hybridized with Bee Colony Optimization is proposed [12] . The method is applied in the medical data set and minimal reduct set is found. The proposed method is compared with Quick Reduct, Entropy based Reduct and with Genetic Algorithm, Particle Swarm Optimization and Ant Colony Optimization hybridized with Rough Set. The solutions provided by Quick Reduct and Entropy based Reduct was close to minimal reduct set. But the solution is not optimal. The GenRSAR, AntRSAR, PSO-RSAR are also performing well but there is no consistency in the result. This is because these algorithms dealt with random parameters. Bee-RSAR exhibits constant and better performance on medical data set.

To offer better conditions on subsequent analysis and to reduce the complexity of data, the data is pre-processed. Rough Set Theory was applied in three pre-processing steps like [16] Discretization, Feature Selection and Instance Selection. Experimental results have shown that the rough Set Theory consider only feature with large dependency. K-NN classifier is used to validate the result.

Similar objects are determined by using Expectation Maximization (EM) Clustering algorithm [17] . The features generated by this method are compared with the Fuzzy Rough Feature Selection and Tolerance based Feature Selection. The important aspect to reduce the irrelevant and redundant features is that they lead to slow learning and low accuracy. Rough set theory used data dependencies to reduce the input dimensionality. Function of the approximation is used to calculate the measure of dependency. The proposed method produces a smaller number of attributes compared with other Feature Selection algorithms. If also improved the average accuracy of classifiers like J48, JRIP and CART.

Quick Reduct [18] Algorithm is used to reduce the number of genes from gene expression data. By using the Quick Reduct Algorithm the minimal data set is obtained. Rough Set Theory is used to pre-process the data to mine suitable rules. For post processing, Formal concept Analysis is used for mining rules. The rules are mined to extract knowledge and the most important factors that affect the decision making [16] . The aim of using Rough Set Theory is that it will help in prediction where as formal Concept Analysis is used in describing the data. This model is believed to be useful for decision making in the medical field.

PSO [24] is used for Feature Reduction. Here, the patients the classified as diseased and non-diseased with the aid of Artificial Neural Network. The parameters such as Regression, Performance plot, Confusion Matrix and ROC Values are used to analyze the performance. The performance of the whole network is increased after applying PSO.

V. Data Set Description

The data set used for experimentation is the bench mark dataset of Heart disease which is obtained from UCI. The data base consists of totally 4 bench mark data set namely Cleveland data set, Hungary data set, Statlog data set and long beach data set. Among all these the Cleveland data set is most used by all researches since the missing data is less when compared with other data sets. The Cleveland data set consists of total of 76 attributes. But all the published articles refer only 14 attributes as are listed below in Table I, where is given the description of the data attributes used for experimentation.

Table I Description of the data set

No	Name	Description
1	Age	Age in Years
2	Sex	1=male, 0=female
3	Cp	Chest pain type (1= typical angina, 2= atypical angina, 3=non-anginal pain, 4= asymptomatic)
4	trestbps	Resting blood sugar (in mmHg on admission to hospital)
5	Chol	Serum cholesterol in mg/dl
6	Fbs	Fasting blood sugar > 120 mg/dl (1=true, 0=false)
7	restecg	Resting electrocardiographic results (0=normal, 1=having ST-T wave abnormality, 2 = left ventricular hypertrophy)

8	thalach	Maximum heart rate achieved
9	exang	Exercise induced angina
10	oldpeak	ST depression induced by exercise relative to rest
11	slope	Slope of the peak exercise ST segment(1=upsloping, 2=flat, 3= downsloping)
12	Ca	Number of major vessels colored by fluoroscopy
13	Thal	3= normal, 6=fixed defect, 7=reversible defect
14	Num	Class (0=healthy, 1=have heart disease)

VI. Proposed Work

The Intelligent Hybrid Quick Reduct Particle Swarm Optimization [IHQRPSO] algorithm combines both the methodology of PSO and QR algorithms to provide some better results. The proposed algorithm function as follows:

The first step is to initialize the Conditional and Decision features (C, D), the swarm size (S), cognitive and social components (c1 and c2), random values (r1 and r2) and inertia (ω). The c1 and c2 are initialized to be 2.0. The random values of r1 and r2 should be between 0 and 1. If the inertia value is greater, then it will result in strong exploration ability and smaller value will result in strong exploitation ability. The inertia value is too assigned between 0 and 1. It is assigned to be 0.33. The swarm size is fixed as 20.

After initializing the variables, the data set is set to be empty. The empty data set is assigned to a variable. After checking the condition $\forall x \in (C - R)$, the Data Fitness value (DFV) is calculated. The DFV is compared with Particle Best. If the Fitness Values is greater the Particle best (P_{best}) then Data Fitness Value is assigned as P_{best} . Again the P_{best} compared with Global Best (G_{best}). If G_{best} is smaller than P_{best} mean, G_{best} is assigned to be the value of P_{best} .

After storing the G_{best} value, the QR Algorithm is used to compare the vales. If $\gamma_{RU}(x) > \gamma_T(D)$, then the empty attribute which is initially assigned will be replaced by the new G_{best} value. G_{best} value is taken only if the above condition is satisfied. Or else, the velocity is again updated by using the equation 12.

$$V_{id} = V_{id} + c1r1(Pid - Xid) + c2r2(Pgd - Xid) \quad \text{-----> 12}$$

$$Xid = Xid + Vid$$

The process goes on repeating until the G_{best} value obtained satisfied the $\gamma_{RU}(x) > \gamma_T(D)$ condition. The condition stops when the classification accuracy of the decisive feature $\gamma_R(D)$ is equal to the classification accuracy of $\gamma_R(C)$. The final output will the optimal data set. The pseudo code of the algorithm is as follows:

Pseudo code: IHQRPSO Algorithm:

Input:

Data set

Algorithm:

Initialize C, D, S, c1, c2, r1, r2, ω

For each particle i in S do

{
 $R \leftarrow \{ \}$

}

do

$T \leftarrow R$

$\forall x \in (C - R)$

 Calculate DFV

 If $DFV > P_{best}$

 Set $P_{best} = DFV$

 If $P_{best} > G_{best}$

$G_{best} = P_{best}$

 If $\gamma_{RU}(x) > \gamma_T(D)$

$T \leftarrow R \cup (X)$

$R \leftarrow T$

 Calculate and update velocity by using the equation12

Until

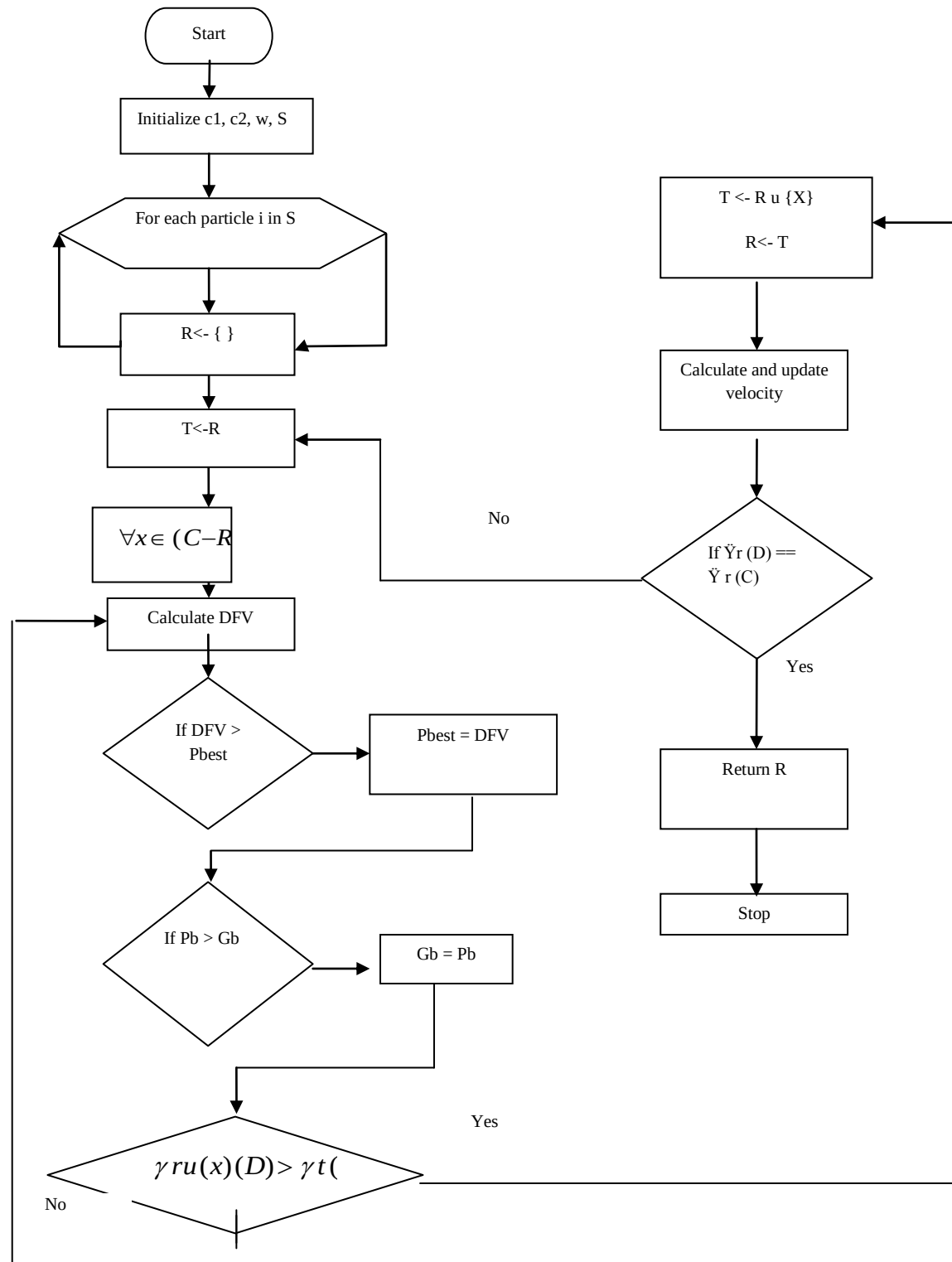
If $\gamma_R(D) == \gamma_C(D)$

Return R

Output:

Optimal Data set

Figure 1: The data flow of IHQRPSO Algorithm



VII. Results and Discussions

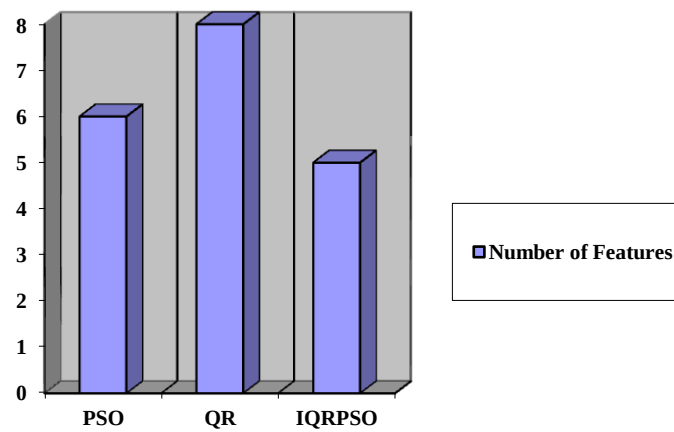
The two algorithms PSO and QR are combined as an IHQRPSO Algorithm. For experimentation, the Matlab is used to simulate the algorithms. It is observed that the proposed algorithm reduce the number of features without affecting the accuracy of the results. The features obtained after implementing the algorithm is as tabulated in Table II.

Table II Features obtained after implementing the algorithms

PSO	QR	IQRPSO
Cp	Cp	Cp
Exang	Exang	Exang
Slope	Slope	Slope
Ca	Ca	Ca
Thal	Thal	Thal
Restecg	Chol	
	Thalach	
	Oldpeak	

The Figure 2 shows the reduced number of attributes by using IHQRPSO Algorithm.

Figure 2: Reduced number of attributes



The performance of all algorithms is calculated by computing its Sensitivity, Specificity and Accuracy. It can be calculated as follows

$$Accuracy = \frac{(TP + FN)}{(TN + TP + FN + FP)} * 100 \quad \text{-----} \rightarrow 13$$

$$Sensitivity = \frac{TP}{TP + FN} * 100 \quad \text{-----} \rightarrow 14$$

$$Specificity = \frac{TN}{TN + FN} * 100 \quad \text{-----} \rightarrow 15$$

Sensitivity is used to correctly classify the individual as diseased and Specificity is to classify the individual as disease free. The True Positive Rate (TPR) and False Positive Rate are also calculated. If the Kappa Statistic value is 0.7 or greater than 0.7, then it is said to good statistic correlation. The correlation is found to be better in the case of high Kappa value.

$$True\ Positive\ Rate = \frac{True\ Positive}{(True\ Positive + False\ Negative)} \quad \text{-----} \rightarrow 16$$

$$False\ Positive\ Rate = \frac{False\ Positive}{(False\ Positive + True\ Negative)} \quad \text{-----} \rightarrow 17$$

$$\text{Precision} = \frac{\text{TruePositive}}{(\text{TruePositive} + \text{FalsePositive})} \text{ ----- } \rightarrow 18$$

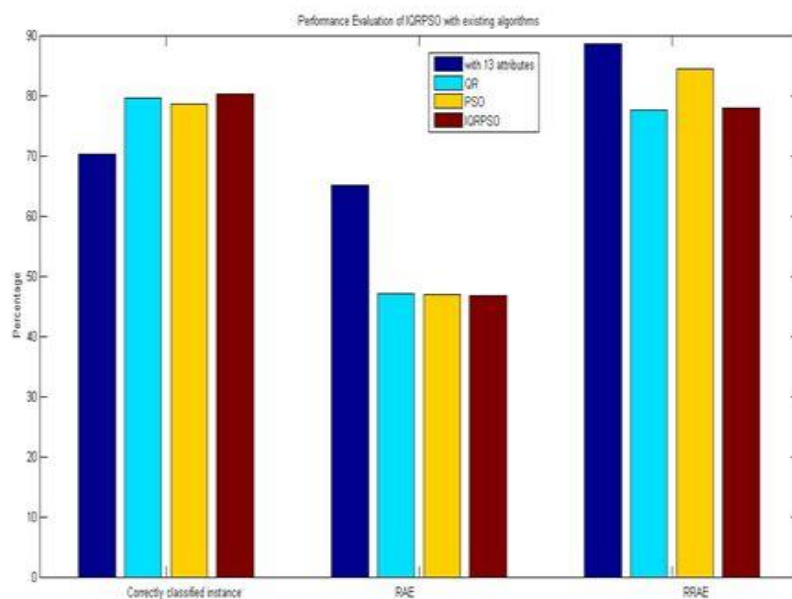
The data set is classified by using Multilayer Perceptron Network. The network is set with 10 hidden neurons and trained using the Levenberg-Marquardt algorithm. The learning rate is set as 0.3 and momentum is 0.2. The Table 3 summarizes the result obtained. The data set is subject to 10 fold cross validation. The cross validation technique is used to estimate the performance of a predictive model. In the 10 fold cross validation; the data sets are divided into 10 sets. Among the 10 data sets, 9 data sets are used for training and 1 is used for testing. The Table III lists out the values obtained after implementing the algorithms.

Table III: Values obtained after implementing the algorithms

	With 13 attributes	PSO	QR	IQRPSO
Correctly classified instance	70.29	79.53	78.54	80.19
Kappa statistic	0.39	0.56	0.58	0.60
MAE	0.32	0.23	0.23	0.23
RMSE	0.44	0.42	0.38	0.38
RAE	65.07	46.98	47.09	46.67
RRAE	88.55	84.42	77.57	77.90
TPR	0.70	0.78	0.79	0.80
FPR	0.31	0.2	0.21	0.2
Precision	0.70	0.78	0.79	0.8
Recall	0.70	0.78	0.79	0.8
ROC Area	0.78	0.84	0.85	0.85

It is observed from the Table 3 that after obtaining the optimal data set by using IHQRPSO algorithm, the Accuracy has been increased. Accuracy level obtained is 80.19% which is higher than other two existing algorithms. The kappa statistic value is also increased. When comparing with other measures like Mean Absolute Error (MAE), Relative Mean Squared Error (RMSE), Relative Absolute Error (RAE), Root Relative Absolute Error (RRAE), True Positive Rate (TPR), False Positive Rate (FPR), Precision, Recall and ROC Area, the optimal data set reduced by using IQRPSO algorithm shows promising results.

Figure 3: Comparison of performance based on Accuracy, RAE and RRAE



The Figure 3 reflects the performance of IHQRPSO with respect to correctly classified instance, RAE and RRAE. It is observed that the classification accuracy is increased with the reduced data set obtained from the

proposed algorithm. Similarly, the RAE and RRAE value is also decreased which supports the performance of proposed IQRPSO in attribute reduction.

Figure 4: Performance analysis of proposed algorithm with existing algorithms

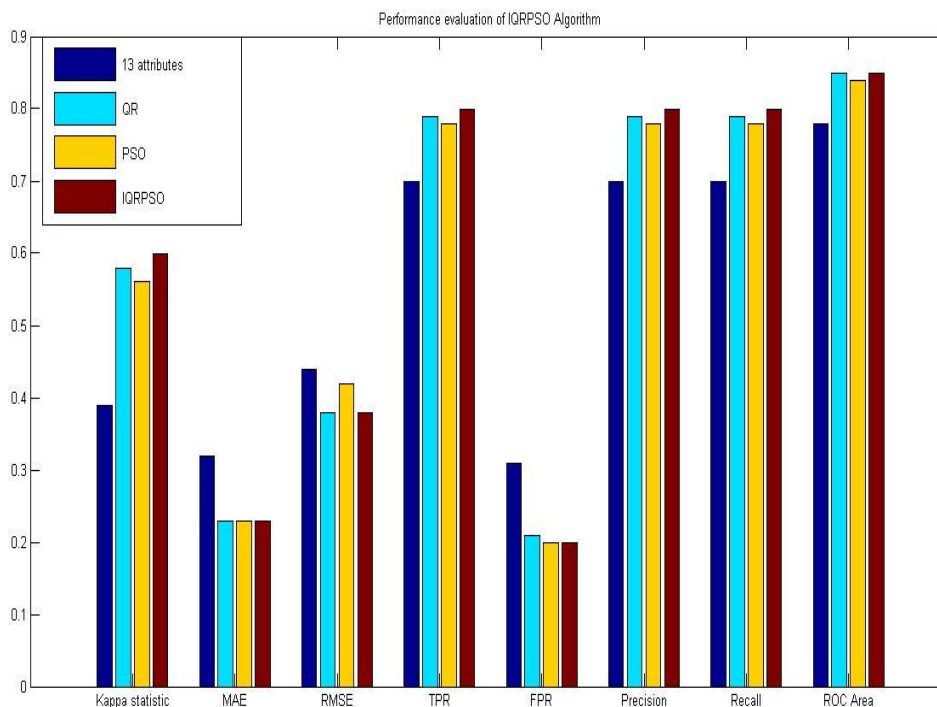
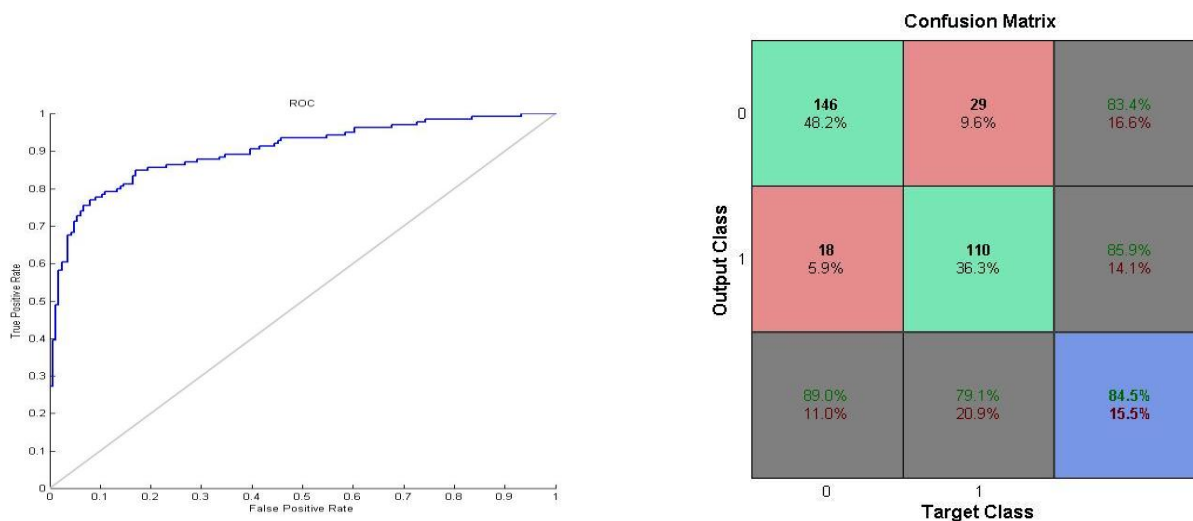


Figure 4 shows the Performance of the algorithm by further analyzing with the parameters like Kappa Statistic, MAE, RMSE, TPR, FPR, Precision, Recall and ROC Area. The graph supports that proposed algorithm can be used for Feature Reduction in Heart Disease data set.

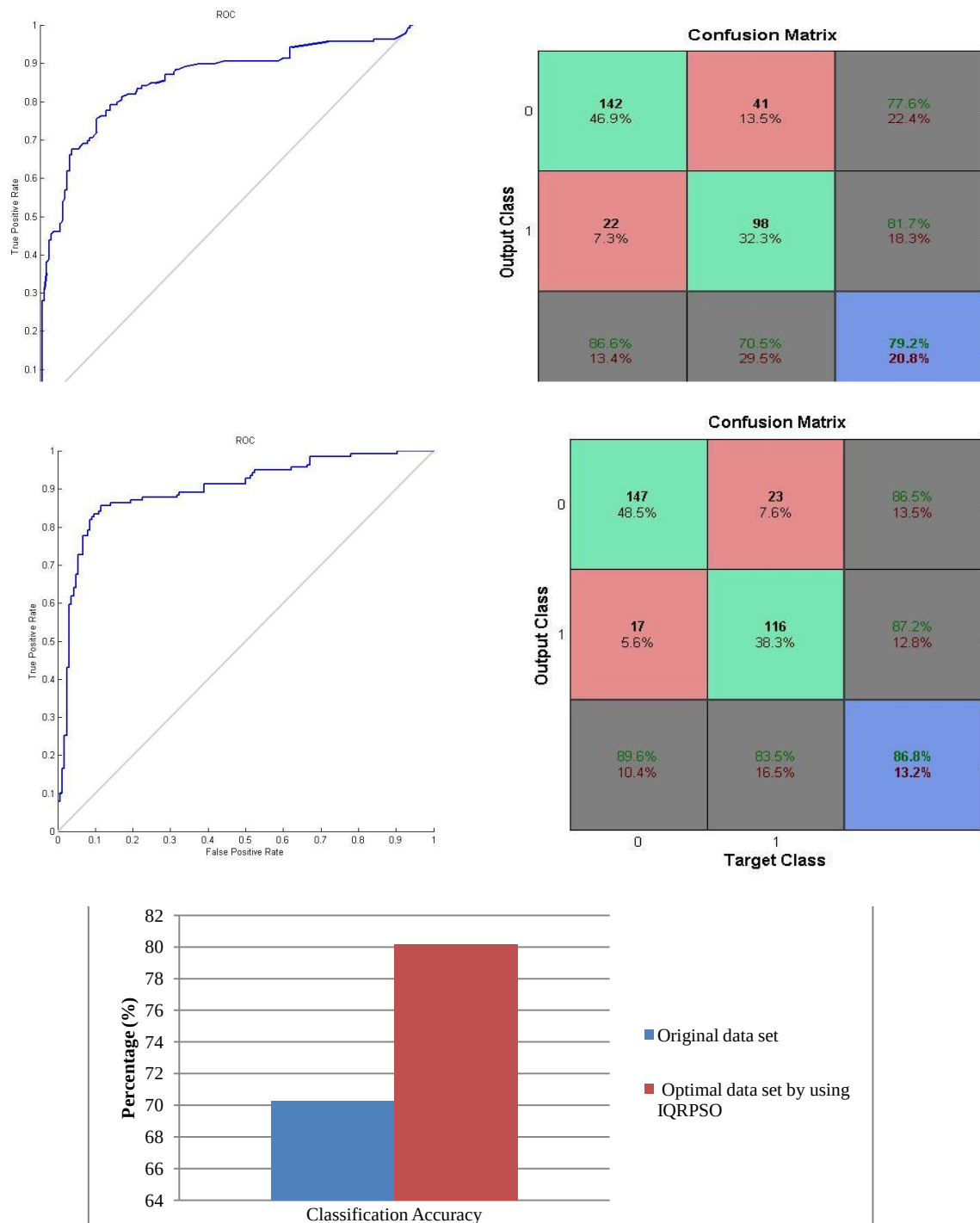
Figure 5: The ROC Curve and Confusion Matrix for QR Algorithm



	Disease Present	Disease Absent	
Test Positive	True Positive	False Positive	Total Positive
Test Negative	False negative	True negative	Total Negative
	Total with disease	Total without disease	Grand Total

ROC curve is the trade-off between the Sensitivity and Specificity. The closer the curve to the left hand border, the more is the accuracy. The Figure 7 give the Confusion Matrix and ROC curve for the PSO Algorithm.

Figure 7: Confusion Matrix and ROC Curve of PSO Algorithm



VIII. Conclusion

Feature reduction is an imperative task in Medical Data Mining. It is a field which requires more accuracy and precise information, since the data used deals with health. Hence, tumbling the attribute number in Medical Data Mining is a crucial part. The optimal data set should not affect the precision of knowledge which is obtained from the data. Different algorithms are applied to obtain the optimal data set by reducing the feature. It is observed that the proposed Intelligent Quick Reduct Particle Swarm Optimization Algorithm is good at producing an optimal data set. The optimal set obtained by using the proposed algorithm does not affect the

accuracy. The accuracy level was increased upto 80% by using this algorithm. The future direction in this research is to test with bigger real time data set.

References

- [1] www.nytimes.com/2009/07/29/opinion/29hall.html.
- [2] Huan Liu and Hiroshi Motoda, Rudy Setiono and Zheng Zhao. "Feature Selection: An Everlasting Frontier in Data Mining", *JMLR: The 4th Workshop on Feature Selection and Data Mining*, 2010.
- [3] World Health Organization. Strategic priorities of the WHO Cardiovascular Disease programme. Available online at URL: <http://www.who.int/whr/2000>. Last accessed February 2006.
- [4] Chen A.H., "HDPS: Heart Disease Prediction System", *Computing in Cardiology*, ISSN 0276-6574, pp 557-560, 2011.
- [5] Latha Parthiban and R.Subramanian, "Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm", *International Journal of Biological and Life Sciences*, Vol 3(3), pp157-160, 2007.
- [6] Elbedwehy M.N, "Detection of Heart Disease using Binary Particle Swarm Optimization", *Proceeding of the Federated Conference in Computer Science and Information System*, 2012, pp 177-182.
- [7] Bing Xue," Multi Objective Particle Swarm Optimization for Feature Selection", *GECCO '12*, 2012.
- [8] Alper unler, Alper Murat and Ratna Babu Chinnam, "m2PSO: A maximum relevance minimum redundancy feature selection method based on Swarm Intelligence for SVM Classification", *Elsevier*, 2011, pp 4625-4641.
- [9] Xiangyang Wang, Jie Yang, Xiaolong Tens and Weijian Xia, Richard Jenson, " Feature selection based on Rough Set and Particle Swarm Optimization", *Pattern Recognition Letters*, 2007, pp: 459-471.
- [10] http://en.wikipedia.org/wiki/Rough_set
- [11] Frida Coaquira and Edgar Acuna, " Application of Rough Set Theory in Data Preprocessing for Knowledge Discovery", *Proceedings of the World Congress on Engineering and Computer Science*, 2007.
- [12] N. Suguna and Dr. K. Thanuskodi, " A Novel Rough Set Reduct Algorithm for Medical Doman based on Bee Colony Optimization", *Journal of Computing*, Vol 2(6), pp 49-54, 2010.
- [13] Yuanning Liu, Gang Wang, Huiling Chen, Hao Dong, Xiaodong Zhu and Sujing Wang " An improved Particle Swarm Optimization for Feature Selection", *Journal of Bionic Engineering*, Vol 8(2), 2011.
- [14] Ismail Babaoglu, Oguz Findik, Erkan Ulker and Nazef Aygul, " A Novel Hybrid Classification Method with PSO and K-nn algorithm for diagnosis of Coronary artery disease using exercise stress test data", *International journal of Innovative Computing*, Volume 8(5), May 2012.
- [15] Benxian Yue, Weihong Yao, Ajith Abraham and Hongbo Liu, " A New Rough Set Reduct Algorithm based on Particle Swarm Optimization", *IWINAC '07*, LNCS 4527, pp 397-409, © Springer Verlag, 2007.
- [16] Mona Nagy Elbedwedhy, Hossam M.Zawbaa, Naveen Ghali and About Ella Hassanien, " Detection of Heart Disease using Binary Particle Swarm Optimization ", *Proceedings of Federated Conference on Computer Science and Information System*, pp 177-182, © IEEE, 2012.
- [17] Faridah Fazayeli, Lipo Wang and Jacek Mandziuk, " Feature Selection based on Rough Set Theory and Expectation Maximization Clustering Algorithm", *RSCTC 2008*, PP 272-282, 2008.
- [18] K. Anitha and Dr. P. Venkatesan, " Feature Selection by Rough-Reduct Algorithm", *IJRSET*, Vol. 2 (8), 2013.
- [19] B.K. Tripathy, D.P. Acharjya and V.Cynthia, " A Framework for Intelligent Medical Diagnosis using Rough Set with Formal Concept Analysis", *International Journal of Artificial Intelligence and Applications*, Vol. 2 (2), 2011.
- [20] Sivagowry.S, Dr. Durairaj. M. (2014). PSO-An Intellectual Technique for Feature Reduction on Heart Malady Anticipation Data. *International journal of Advanced Research in Computer Science and Software Engineering*, 4(9), 610-621.
- [21] Sivagowry.S, Dr. Durairaj.M. "A Pragmatic Approach of Preprocessing the Data Set for Heart Disease Prediction." *International Journal of Innovative Research in Computer and Communication Engineering* 2.11 (2014): 6457-6465.
- [22] Sivagowry, S., M. Durairaj, and A. Persia. "An empirical study on applying data mining techniques for the analysis and prediction of heart disease." *Information Communication and Embedded Systems (ICICES)*, 2013 International Conference on. IEEE, 2013.
- [23] Durairaj.M, Sivagowry.S. "Feature Diminution by Using Particle Swarm Optimization for Envisaging the Heart Syndrome." *International Journal of Information Technology and Computer Science(IJITCS)* (7).2 (2015): 35-43.
- [24] Durairaj.M, Sivagowry.S. "A Survey on Particle Swarm Optimization and Rough Set Theory in Feature Selection for Heart Disease Prediction." *International Journal of Computer Science and Mobile Computing (IJCSMC)* Vol 4 (3) (2015): 87-92.