A Rough Set Pooled Fitness Function Based Particle Swarm Optimization Algorithm using Golden Ratio Principle for Feature Selection

K. Saravanapriya, J. Bagyamani

Abstract*: *Particle Swarm Optimization, a nature based stochastic evolutionary algorithm that iteratively tries to improvise the solution pertaining to a particular objective function. The problem becomes challenging if the objective function is not properly identified nor it is properly been evaluated which results in slow convergence and inability to find the optimal solution. Hence, we propose a novel rough set based particle swarm optimization algorithm using golden ratio principle for an efficient feature selection process that focusses on two objectives: First, that results in a reduced subset of features without conceding the originality of the data and the second is that yields a high optimal result. Since many subset of features might result with a meaningful solution, we have used the golden ratio principle to extract the most reduced subset with a high optimal solution. The algorithm has been tested over several benchmark datasets. The results shows that the proposed algorithm identifies a small set of features without convincing the optimal solution, thus able to achieve the stated objectives.

Keywords* : · *Fitness Function · Classification · Decision Tree · Feature Selection · Golden Ratio Principle · Particle Swarm Optimization · Quick Reduct Algorithm · Rough set · Support Vector Machine · Naïve Bayes.

I. INTRODUCTION

Feature selection is a data preprocessing system that is widely used for classification problems. It aims to identify a subset of unique features without compromising the accuracy and the efficiency of the classification technique. Most high dimensional data possess least the number of instances yet greater the number of features [1]. Most of the features are either redundant or of no relevance. These may be difficult to evaluate. Rough set theory is extensively used to rectify the persisting problem of feature selection proposed by Pawlak in 1982 [1], has been extensively used in machine learning and data mining to identify the data dependencies and identify the minimal subset of the features with the same discernibility as that of the entire features present in the dataset [2]. Based on the searching criteria, feature selection remains tedious due to the exponential increase of search space based on the number of available features [3]. Particles Swarm Optimization (PSO) is preferred for its simplicity, ease of implementation and economic feasibility. Here, we propose a fitness function combined rough Set based particle swarm optimization algorithm (RSGRPSO) for an efficient feature selection process. It focusses on two objectives: (1) the best fitness solution (to be maximized), and (2) the number of features (to be minimized). This methodology is similar to the two-part objective functions which has a goodness-of-fit term and a regularization term [3], specifically while the effect of the regularization term is to reduce the parameter space.

Further the paper proceeds in the following manner: Section 2 that describes about some basic concepts of Classification, Particle swarm optimization, Rough Sets, and about the related works. The proposed MORSPSO algorithm is demonstrated in Section 3. The performance accuracy of the proposed algorithm and its comparison with the QRA is depicted in Section 4 and finally the conclusion, findings and future scope is elicited in Section 5.

II. PRELIMINARIES AND RELATED WORKS

A. Rough Set Theory

Rough Sets (RST) [7] discovers the structural relationship from the noisy data, it uses the discretization technique on discrete-valued attributes and continues-valued attribute [8]. It has been extensively used for feature selection by most of the authors. Velayutham proposed an improved rough set algorithm for finding optimal attribute reduct. They have introduced the usage of threshold value in the traditional Quick Reduct method for finding the optimal reduct [9]. Thangavel has proposed a modified QRA that discards the objects that took part in building the lower approximation that in turn decreases the horizontal information size The author also discusses the performance of various reduct algorithms for the construction of efficient rules [10].

Let I = {U, MCA∪{NDA}} or I = (U, NDA) is the Information system, U defines the non-empty, universal, finite set of instances, MCA and NDA denotes the Conditional (CA) and Decision (DA) attribute [11].

$\forall c \in$ MCA, there is an equivalent function, where Vc represent c's set of values.

$IND\,(MCA) = \{(a, b) \in U \times U / \forall c \in MCA, fc(a) = fc(b)\}$ (1)

Revised Manuscript Received on October 15, 2019.

K. Saravanapriya, Research Scholar, Department of Compter Science, Periyar University, Salem, India. Email: rajpriya2109@gmail.com

J. Bagyamani, Associate Professor, Department of Computer Science, Government Arts College , Dharmapuri, Dharmapuri, India. Email: bagya.gac@gmail.com

IND (MCA) generates the partition of U which is represented as U/MCA. If (a, b) ∊ IND (MCA), then a, b is said to be indiscernible by attributes MCA. MCA is the notational representation for the equivalence class of M-indiscernibility relation. Let, A⊆U, MCA_L(A) is the lower approximation and MCA_U(A) is the upper approximation can be denoted as:

$$MCA\_L(A) = \{a \in U \mid [a]MCA \subseteq A\} \quad (2)$$

$$MCA\_U(A) = \{A \in U \mid [A]\, MCA \cap A \;\; f = \emptyset \subseteq A\} \quad (3)$$

Let MCA and NCA ⊆ C be the relation equivalent to U, so, the boundary region (BND), positive_region (POS) and negative_region (NEG) is:

$$POS_{MCA}(NDA) = U - \bigcup_{a \in \frac{U}{NDA}} MCA\_L(A) \quad (4)$$

$$NEG_{MCA}(NDA) = U - \bigcup_{a \in \frac{U}{NDA}} MCA\_U(A) \quad (5)$$

$$BND_{MCA}(NDA) = \bigcup_{a \in \frac{U}{NDA}} MCA\_L(A) - \bigcup_{a \in \frac{U}{NDA}} MCA\_U(A) \quad (6)$$

NDA is dependent on MCA in a degree g (0≤g≤1) is denoted by

$$g = \gamma_{MCA}(NDA) = \frac{|POS_{MCA}(NDA)|}{|U|} \quad (7)$$

Where $\gamma_{MCA}(NDA)$ is the classification quality [12]. If g=1, NDA completely dependent on MCA; if 0<g<1, NDA partly depends on MCA, and if g=0, then NDA is independent of MCA.

Attribute reduct aims at removing the redundant attributes without compromising the originality of the dataset [13]. The reduced set of attributes is given as:

$$Reduct = \{R \subseteq MCA \,/\, \gamma_R(NDA) = \gamma_{MCA}(NDA),\ \forall B \subset R,\ \gamma_B(NDA) \neq \gamma_C(NDA)\} \quad (8)$$

A dataset may contain any number of reduced attribute sets. The optimal subset is given as:

$$Reduct_{min} = \{R \in Reduct / \forall R^j \in Reduct, /R/ \leq /R^j/ \quad (9)$$

B. Quick Reduct Algorithm

Quick Reduct Algorithm (QRA) [26] is used to derive a reduced set of attribute. We compare the equivalence relations that results from a set of attributes, the resultant subset gives the similar prognostic ability as that of the source dataset [11]. The QRA Algorithm is represented in Fig 1.

C. Particle Swarm Optimization

Particle Swarm Optimization (PSO) depicts the behavior of swarm of Birds. Though there exists many evolutionary algorithms, PSO is widely used by the researchers for its ease of use, computationally less expensive, fewer parameter requirement, and the one which converges quickly. Particle Swarm Optimization has two key steps: 1. Representation of particles, 2. Fitness function. For C particles and an S-dimensional search space, let Pop represent the random population, V, the velocity. V and P are given by S X C matrix [2].

```
Input
CF - Conditional features.
DF - Decision features.
Output
Reduct Red
Red ← {}
Repeat
    Temp←Red
    ∀a∈(CF-Red)
    If γRed∪{a}(DF) > γTemp(DF)
    Temp←Red{a}
    Red←Temp
While γRed(DF) = γCF(DF)
Return Red
```

Fig. 1. Quick Reduct Algorithm

Per_best represents the current personal best information exhibited so far. Eq. (16) helps the particle return to the last position that is better than the current one. Gl_best is the current global best solution obtained by the entire particle population so far and is represented as 1 x S vector. S dimension takes thrice the value of the number of instances of the dataset or it can be determined based on the problem size.

The initialization of particles is done randomly in the range of [0, 1]. The initialization of Velocity V is based on the initial particle position. W in Eq.(14) represents the inertia weight that helps the particle move with the same velocity and in the same side. The learning factors, L1 and L2 is often assigned 2 but can vary with a range of [0, 4].

$$V_{i,j}(z+1) = W + Pop_{i,j}(z) + L1 * r_{i,j}(1) * (Per\_best_{i,j}(z) - Pop_{i,j}(z)) + L2 * r_{i,j}(2) * (Gl\_best_j(z) - Pop_{i,j}(z)) \quad (10)$$

$$Pop_{i,j}(z+1) = Pop_{i,j}(z) + V_{i,j}(z+1) \quad (11)$$

$$Per_{best\,i,j} = \begin{cases} Pop_{i,j}(z+1), \text{if } fitness\,(x_i(z+1)) \text{ is better than } Per_{best\,i,j}(z) \\ Per_{best\,i}(z) \quad \text{otherwise} \end{cases} \quad (12)$$

PSO exhibits two problems: Firstly, the Global version, is fast, yet for some problems, it converges to the local optimum. Secondly, the Local version which is slow, escapes to be easily stuck within local optimum.

D. Related works

Feature Selection algorithm is categorized into Wrapper and Filter methods. Filter Method is free of any learning model and it suits for high dimensional data analysis. It is fast and computationally simple. Whereas the wrapper method evaluates the features by means of a learning model. Its time complexity is large compared to filter method. The details of the works carried out is depicted in Table I.

III. PROPOSED WORK

The proposed work focusses on achieving the objectives of the feature selection that is to

Table- I: Related Works

Papers	FS Searching Technique	Classification Techniques	Results
Mateusz Adam-czyk[19]	Asynchronous PSO	-	Since the data is frequently updated and the algorithm is checked for convergence, if needed, the global best position is updated. The particles act dynamically to find the best solution.
Pedram Ghamisi [20]	Hybrid PSO and GA	SVM	It helps to identify roads within a stipulated CPU processing time. It also strives to improve the classification accuracy of the data.
Hannah Inbarani {15]	Hybrid PSO + rough sets	Naïve Bayes, Bayes Net, KStar	Aid in the dimensionality reduction, and consumes less time.
Binh Tran [17]	Bare-bone particle swarm optimization (BBPSO) for discretization	-	Effectively discretizes multiple features and improves the classification performance.
Liam Cervante et al [25]	Binary PSO + Mutual Information and BinaryPSO + Entropy based	Decision trees	Selects smaller number of feature with good classification accuracy. The algorithms were applied with different weights for relevance and redundancy.
Bing Xue [24]	PSO based Single Objective and Multi Objective approaches, Mutual Information, Entropy, RS and Probability Rough Set	SVM, KNN, DT, NB	Multi-objective algorithms performs better than the single objective algorithms,
Xiang yang Wang et al [16]	PSORSFS, POSAR, CEAR, DISMAR, GAAR	-	It could be able to find optimal values on certain datasets. It also tells that PSO outperforms than GA.
Rafael [14]	PSO based on Rough Set theory		Splits the particles into two stages. One with used and unused features and the resultant swarm is given as the input for the second stage.

minimize the number of features selected without affecting the classification accuracy of the dataset. We propose a novel fitness function that results with a high optimal value without affecting the classification accuracy. With the intention of increasing the optimal value, we include the golden ratio principle whose inclusion yields a best fitness solution. The explanation for golden ratio is as follows.

A. Golden Ratio Principle

The Golden ratio is a mathematical concept which has been widely used in most of the scientific theories. When a line is divided into smaller and larger part, the smaller part divides the larger part which is equal to the entire length divided by the larger part. It is otherwise called as the divine proportion, medial section, golden section, golden mean etc. It is denoted by Φ [21]. The golden ratio is diagrammatically depicted in Fig 2.

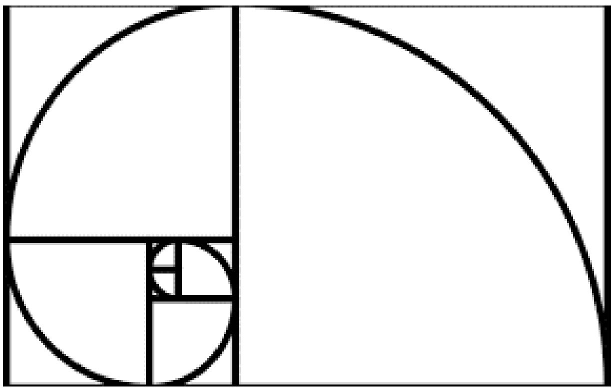


Fig. 2. Golden ratio

Algebraically, it is written for the quantities h and g with h > g > 0 as:

$$(h+g)/h = h/g = \Phi \quad (13)$$

Its value raise up to ∞, so it is rounded to 1.618. Here in this paper, we have used this golden ratio principle for calculating the Objective function.

$$\textit{Objective Function} = (\textit{Gamma} - 1 + \textit{Count of Selected Attributes})/1.618 \quad (14)$$

For Example, if the number of minimum selected attributes for the above mentioned Toy dataset in Table 1 are 3, 4, 5.

Objective Function = (1 – 1 + 3) / 1.618 = 1.8541

The results shows that with the inclusion of the golden ratio, we were able to identify a reduced subset of features with a global optimum solution

B. Proposed Algorithm

1. Start
2. Generate the random population
3. Initialize
   a. Inertia weight W to 0.8.
   b. Learning Factors L1, L2 to 2.
   c. Gl_best to the first value of X.

d. Per_best is assigned with value X.

e. Fitness is assigned to 0.

f. Velocity V is represented as X*1.0.

4. Repeat Step 4 to 8 until a predetermined condition has been met.
5. Get Reduct
6. Get Reduct (Xi) using Quick reduct algorithm as mentioned in Fig. 1.
7. Update Pbest using Eq.8)
8. Gbest = Bestfit (any Pbest)
9. Update Velocity using Eq.10
10. Update Particle using Eq.11
11. End

IV. EXPERIMENTAL ANALYSIS

The proposed work is applied on six mostly used datasets from the UCI Machine Learning Repository [22] using MATLAB version 2016. The details of the datasets used for the analysis is represented as in Table 3. Here, we have focused mainly on the biological data sets. The Iris dataset is a multivariate dataset that consists of 5 attributes with 150 Instances which includes 3 Class values. The Liver Disorder dataset is a multivariate dataset that has 7 attributes which includes the class value with 345 instances. This is a collection of male dataset from the BUPA Medical Research Limited. The next dataset focusses on the Indian Liver Patients dataset which is also a multivariate dataset that includes 10 attributes and 583 instances out of which 416 datasets were of patients' whose liver has been affected and the remaining were of healthy patients' records. Hepatitis dataset is a multivariate dataset that has 155 instances with 19 attributes. The class value suggests that whether the patient is alive or not. The Lymphographic dataset is a multivariate dataset that includes categorical values with 148 instances and 19 attributes. The class value divides the dataset into 4 categories such as normal find, metastasis, malign lymph, and fibrosis. The PIMA Indian diabetes data refers to the females of ages from 21 to 81. The data set has 768 instances with 9 input attributes including 2 class attributes that enables to predict the status of the patients to be healthy or diabetic [22]. The dataset details has been depicted in Table II.

Table – II. Details of the Datasets Used

S. No	Datasets Used	No. of Attributes	No. of Class Attributes	No. of Instances
1	Iris	4	1	150
2	Liver Disorder	6	1	345
3	ILPD	10	1	583
4	Hepatitis	19	1	155
5	Lymphography	18	1	148
6	Diabetes	8	1	768

The datasets have been tested with the QRA and the proposed RSGRPSO algorithm and the results such as dependency ratio (DR), selected Attributes (SA), best fit value have been compared and is depicted in Table III. For QRA, the Iris dataset resulted with 2 reducts with a dependency ratio of 0.8533, the liver disorder dataset with 5 reducts with 0.9246 as the dependency ratio. The ILPD dataset resulted with a dependency ratio of 0.9177 with 7 subsets. The hepatitis dataset yielded 0.9274 with 6 subsets. The lymphography with 0.8784 with 8 subset. The diabetes dataset resulted with the dependency ratio of 0.8932 with 7 reducts.

As per the RSGRPSO algorithm, the objective function focusses on the set of attributes which has the maximum dependency value along with the inclusion of golden ratio principle as represented in Eq. (15). Though there were specific varied dependency values, we focus only on the maximum of those values for a dataset and, as stated, all the above mentioned dataset resulted with a maximum dependency value 1.0000, but with a different reduced set of attributes. For the Iris dataset, the best fit values obtained were 1.2361 with 1 reduct, 1.8541 with 2 reducts and 2.4722 with 3 reducts. The liver disorder data resulted with a ratio of 1.8541 with 2 attributes, 4.3262 with 6 attributes. The ILPD results with 2.4722 with 3 reducts, and 3.0902 with 4 reducts and 4.3263 with 6 reducts. Hepatitis with 5 reducts with the bestfit of 3.7083, 4.3263 with 6 reducts and 4.9444 with 7 reducts. Lymphography leads to a bestfit value of 2.4722 with 3 attributes, 4.3262 with 6 reducts and 4.9444 with 7 reducts. Diabetes resulted with 1 attributes with a bestfit of 1.2361, 1.8541 with 2 attributes, 2.4722 with 3 reducts, 3.0902 with 4 reducts and 3.7083 with 5 reducts.

Table – III. Selected Attributes and the Best Fit

S.No	Datasets	Max. Dependency value	Best Fit	Particle Count of Best Fit	Selected Attribute
1	Iris	1.0000	1.2361	1	3
2	Liver Disorder	1.0000	1.8541	2	1, 4
3	ILPD	1.0000	2.4722	3	1, 2, 4
4	Hepatitis	1.0000	3.7083	5	1, 2, 3, 14, 17
5	Lmphography	1.0000	2.4722	5	3, 8, 11, 15, 17
6	Diabetes	1.0000	3.7083	5	2, 4, 5, 6, 8

Table IV depicts the best fit and the resultant subset of features for each dataset using RSGRPSO. Though the results of RSGRPSO has minimum reduct with a high dependency ratio, some of those could not achieve better accuracy. So, here the subsets have been chosen based on the maximum accuracy obtained for various classifiers. It also represents the maximum dependency value which plays a vital role on classification accuracy and the minimal set of reduct for each dataset.

It has been observed that For the Iris dataset, the best fit value was 1.2361 with 1 reduct. The liver disorder data resulted with the best fit value of 1.854 with 2 attributes. The ILPD with a minimum value of 2.4722 with 3 reducts. Hepatitis with 5 reducts

Table- IV. Comparison of Dependency Values and the Reduced Set of Attributes of Quick Reduct and RSGRPSO

Datasets	Quick Reduct			RSGRPSO			
	DR	No. of SA	SA	DR	Best fit	No. of SA	SA
Iris	0.8533	2	3, 4	1.0000	**1.2361** 1.8541 2.4722	**1** 2 3	**3** 2, 3 1, 3, 4
Liver Disorder	0.9246	5	2, 3, 4, 5, 6	1.0000	**1.8541** 4.326 2	**2** 6	**1, 4** 1, 2, 3, 4, 5, 6
ILPD	0.9177	7	1, 2, 3, 5, 8, 9, 10	1.0000	**2.4722** 3.0902 4.3263	**3** 4 6	**1, 2, 4** 1, 2, 3, 5 1, 2, 5, 6, 7, 9
Hepatitis	0.9274	6	2, 7, 12, 16, 17, 18	1.0000	**3.7083** 4.3263 4.9444	**5** 6 7	**1, 2, 3, 14, 17** 4,9,10,11,15, 17 2,5,10,12,15, 16, 17
Lymphography	0.8784	8	2, 6, 11, 12, 13, 14, 15, 17	1.0000	2.4722 **3.7083** 4.3263 4.9444	3 **5** 6 7	11,15,17 **3, 8, 11, 15,17** 1,2 5,11,17,18 1,4,10,11,15, 16, 17
Diabetes	0.8932	7	1, 2, 3, 4, 5, 7, 8	1.000	1.8541 2.4722 3.0902 **3.7083**	2 3 4 **5**	1, 3 4, 5, 8 2, 5, 7, 8 **2, 4, 5, 6, 8**

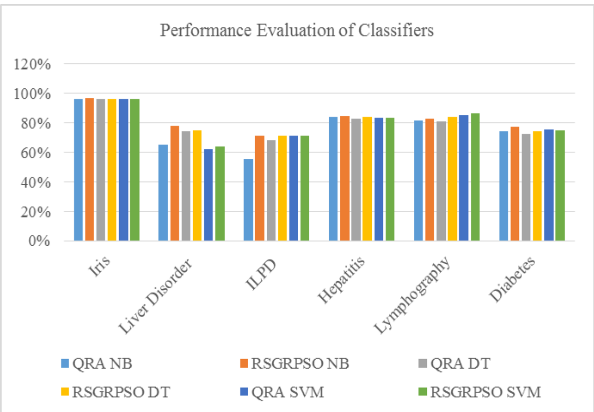
with a best fit value of 3.7083. Lymphography leads to a value of 2.4722 with 5 attributes. Diabetes resulted with 5 attributes with the best fit value of 3.7083.

The dataset with the above said reduced set of features for both QRA and the proposed RSGRPSO algorithm have been applied on well-known classifiers and their performance have been compared and analyzed. The results obtained is shown in Table 6.

Table – V. Classification based on Selected Attributes

Datasets	Quick Reduct Algorithm (%)			RSGRPSO Algorithm (%)		
	NB	DT	SVM	NB	DT	SVM
Iris	96	96	94.6	**96.6**	96	96
Liver Disorder	65	74	62.39	**78**	75	64.10
ILPD	55.7	68	71.3	**71.35**	**71.35**	**71.35**
Hepatitis	84	83	83.2	**84.5**	83.8.	83.2
Lymph.	81.7	81.1	85.1	83	83.7	**86.48**
Diabetes	74	72.3	75.7	77	74	**74.8**

Fig. 3. Performance of Various Classifiers



From the table, we can observe that the proposed RSGRPSO algorithm outperformed for all the dataset. For Iris dataset, the Naïve Bayes classifier resulted with 96.6%. For the liver disorder dataset, Naïve Bayes yielded 78%, similarly, for ILPD, all the three classification algorithm resulted with a similar accuracy rate with 71.35%. For Hepatitis dataset, Naïve Bayes resulted with 84.5%. The Lymphography dataset resulted with 86.48% for SVM. For diabetes dataset, Naïve Bayes gave 77% which is higher than all other algorithms. The maximum accuracy values have been highlighted.

V. CONCLUSION

The proposed algorithm aims to find the optimal solution with a reduced set of attributes such that it satisfies the main goal of feature selection. It selects a random set of population from which it evaluates the objective function that uses the principle of golden ratio, to reach out the optimal solution without compromising the originality of the data sets thus satisfies the objective function. The proposed RSGRPSO converges quickly and could be able to produce the optimum result. The resultant reduced set of attributes have been applied in well-known classification algorithms to evaluate the performance. Of the three algorithms, RSGRPSO based Naïve Bayes classification algorithm outperformed the other two classifiers. In future, the algorithm shall be applied on real time medical data sets so as to help the medical practitioners to diagnose the disease at the earliest and make novel predictions over the diseases.

REFERENCES

1. Zdzisław Pawlak, "Rough Sets", International Journal of Computer and Information Sciences, vol. 11, Issue 5, pp. 341 – 356, 1982.
2. Ramesh Kumar Huda, Haider Banka, "Efficient feature selection and classification algorithm based on PSO and rough sets. Neural Computing and Applications, pp. 1 – 17, 2013. https://doi.org/10.1007/s00521-017-3317-9.
3. Guyon I, Elisseeff A, "An Introduction to variable and feature selection", J Mach Learn Res, Vol. 3, pp:1157–1182, 2003.
4. Dash M, Liu H, "Feature selection for classification", Intell Data Anal, Vol. 1(3), pp: 131–156, 1997.
5. Kalyanmoy Deb, "Multi-Objective Optimization Using Evolutionary Algorithms: An Introduction", Multi-objective Evolutionary Optimisation for Product Design and Manufacturing. Springer, London. https://doi.org/10.1007/978-0-85729-652-8_1.

6. K. Saravanapriya and J. Bagyamani, "Performance Analysis of Classification Algorithms on Diabetes Dataset", International Journal of Computer Sciences and Engineering, Vol 5, Issue 9, pp: 15 – 20, 2017.
7. Cheng Li, Zhihuan Song, Ping Li (2004) Bearing fault detection via wavelet packet transform and rough set theory. Fifth World Congress on Intelligent Control and Automation. IEEE Explore China. https://doi.org/10.1109/WCICA.2004.1340953.
8. Zdzisław Pawlak, "Rough Set Theory and its applications", Journal of Telecommunications and Information Technology, Vol. 3: pp. 7 – 10, 2002.
9. M C. Velayutham and K. Thangavel, "Improved Rough Set Algorithms for Optimal Attribute Reduct", Journal of Electronic Science and Technology, 9(2), pp. 108 – 117, 2011.
10. Thangavel. K, Peethalakshmi and P. Jaganathan, "A Comparative Analysis of Feature Selection Algorithms Based on Rough set Theory", International Journal of Soft Computing, Vol. 1, pp. 288 – 294, 2006,
11. K. Saravanapriya, S. Ramadass and J. Bagyamani, "Performance Analysis of Quick Reduct and Multiple Optimal Reduct Algorithm Based on Dependency Ratio for Feature Selection", Proceedings in International Conference on Electrical and Electronics Engineering, 978-1-5386-4304-4/18 © IEEE, 2018.
12. Z. Pawlak, "Rough set approach to knowledge-based decision support". European Journal of Operational Research, 99(1), pp: 48 – 57, 1997.
13. Javad rahimpoor anaraki, Mahdi Eftekhari, "Rough set based feature selection: A Review", IEEE, Information and Knowledge Technology (IKT), 2013, https://doi.org/10.1109/IKT.2013.6620083.
14. Rafael Bello, Yudel Gomez, Ann Nowe, María M. García, "Two Step Particle Swarm Optimization to Solve the Feature Selection Problem", Seventh International Conference on Intelligent Systems Design and Applications, IEEE 2007, DOI: 10.1109/ISDA.2007.101.
15. H. Hannah Inbarani, Ahmad Taher Azar, G. Jothi, "Supervised hybrid feature selection based on PSO and rough sets for medical diagnosis", Computer Methods and Programs in Biomedicine, Elsevier, Vol. 113, Issue. 1, pp: 175 – 185, 2014. https://doi.org/10.1016/j.cmpb.2013.10.007.
16. Xiangyang Wang, Jie Yang, Xiaolong Teng, Weijun Xia, Richard Jensen, "Feature Selection based on Rough Sets and Particle Swarm Optimization", Elsevier Science Inc. New York, 28(4), pp. 459 – 471, 2007. https://doi.org/10.1016/j.patrec.2006.09.2003.
17. Binh Tran, Bing Xue, and Mengjie Zhang, "Bare-Bone Particle Swarm Optimisation for Simultaneously Discretising and Selecting Features for High-Dimensional Classification", EvoApplications, pp. 701 – 718, 2016, https://doi.org/10.1007/978-3-319-31204-0_45.
18. Chakraborty, Bishwajit, "Feature Subset Selection by Particle Swarm Optimization with Fuzzy Fitness Function", IEEE Xplore, 2008, https://doi.org/10.1109/ISKE.2008.4731082.
19. Mateusz Adamczyk, "Parallel Feature Selection Algorithm based on Rough Sets and Particle Swarm Optimization", IEEE Proceedings of the 2014 Federated Conference on Computer Science and Information Systems. ACSIS. 2, pp. 43–50.
20. Pedram Ghamisi, "Feature Selection Based on Hybridization of Genetic Algorithm and Particle Swarm Optimization", IEEE Geoscience and Remote Sensing Letters. Vol. 12(2), 2015, pp: 309-313.
21. Wikipedia - https://en.wikipedia.org/wiki/Golden_ratio
22. UCI machine Learning Repository- http://www.archive.ics.uci.edu.
23. https://en.wikipedia.org/wiki/Cohen%27s_kappa.
24. Bing Xue, Mengjie Zhang and Will N. Browne, "A Comprehensive Comparison on Evolutionary Feature Selection Approaches to Classification", IEEE, International Journal of Computational Intelligence and Applications, Vol. 14 (2), 2015 https://doi.org/10.1142/S146902681550008X.
25. Liam Cervante, Bing Xue, Mengjie Zhang, Lin Shang, "Binary Particle Swarm Optimisation for feature Selection Approaches to Classification", IEEE, International Journal of Computational Intelligence and Applications, Vol. 14(2), pp. 1656 – 1671, 2015, https://doi.org/10.1109/TSMCB.2012.2227469.
26. K. Thangavel, C. Velayutham, "Unsupervised feature selection in digital mammogram image using rough set theory", International Journal of Bioinformatics Research and Applications (IJBRA), Vol. 8(5-6), pp: 436-454, https://doi.org/10.1504/IJBRA.2012.049626.

AUTHORS PROFILE

Ms. K. Saravanapriya is currently pursuing her Ph.D. degree under the guidance of Dr. J. Bagyamani.in the Department of Computer Science, in Periyar University, Salem, India. Her research interests include Data Mining, Computer Graphics. She has 10 years of teaching experience and 4 years of Research experience.

Dr. J. Bagyamani is an Associate Professor and Head in the Department of Computer Applications, in Government Arts College, Dharmapuri, Tamil Nadu, India. She has got 24 years of Teaching Experience and 10 years of Research experience. She has published 15 papers in reputed International Journals, and presented 8 Papers in National and International Conferences and Seminars. She has received Two Best paper awards in International Conferences. She has guided 18 M. Phil Research scholars. Her areas of research and interests include Data mining, Bi-clustering of Gene Expression Data using Heuristic and Meta-heuristic techniques, Optimization algorithms, Web mining and Image Processing.