Journal of Theoretical and Applied Information Technology
31st August 2019. Vol.97. No 16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ANALYSIS OF HYBRID METAHEURISTIC TECHNIQUES ON CLASSIFICATION

ANJU BALA[1], PRITI[2]
[1]Research Scholar, DCSA, M.D.U. Rohtak, Haryana (India)
anjusangwan1988@gmail.com
[2]Assistant Professor, DCSA, M.D.U. Rohtak, Haryana (India)
pritish80@yahoo.co.in

ABSTRACT

This paper discusses different hybrid version of Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Differential Evaluation (DE) and Genetic Algorithm (GA) and analyzes the performance of different hybrid algorithms in terms of classification accuracy. The hybridization is done to remove the limitations of each individual technique by incorporating the advantages of other techniques resulting in better convergence towards global optima. The paper covers various hybrid meta-heuristic techniques including ACO-PSO, ACO-GA, PSO-GA, GA-DE, and ACO-PSO-GA. The analysis has been done on different datasets downloaded from UCI repository using the parameters classification accuracy, sensitivity and specificity. The analysis clearly shows the impact of the hybridization on the classification in terms of accuracy as well as sensitivity and specificity.

KEYWORDS: ACO, PSO, GA, DE, Hybrid, Classification, Meta-heuristic

1. INTRODUCTION

Classification is one of the basic tasks which is becoming critical with technological advancement [1]. It is a process to categories the entities based on their properties like filtering of e-mails to personal, social or promotions based on the content of the e-mail. A number of authors have worked on different algorithms for classification. The classification accuracy is the key to measure the performance of the algorithm along with the running time especially on the huge datasets available today. The classification accuracy, as well as the running time to classify the data, depends upon the quality of the patterns extracted [2]. The quality of raw patterns can be improved by using feature selection as the pre-processing step to remove the redundant variables. Feature selection methods are of three types i.e. wrapper, filter and hybrid based on their techniques. The wrapper technique uses the classifier to determine the significance of the feature while the filter technique uses statistical measure for the same. The wrapper technique is more accurate as compared to filter technique while filter technique is faster as compared to the wrapper technique. The hybrid techniques combine the feature of both and ensembles the classifier within the technique [3]. This paper focuses on the wrapper techniques due to their enhanced accuracy. Different authors proposed various techniques for the feature selection and analyzed their performance which shows that meta-heuristic techniques exhibit better performance as compared to the traditional techniques.

Feature Selection can be considered as an optimization problem which optimizes the accuracy and running time [4]. Optimization problems are complex problems which can be effectively solved by using meta-heuristic techniques. The effective solution is achieved by a combination of exploration and exploitation search used by the meta-heuristic techniques. Various meta-heuristic techniques are proposed in the literature that mimics the behavior of any natural species or process [5]. The capability of such techniques to avoid the local optima depends upon balancing between exploration and exploitation search according to the search space[6]. That’s why this paper applies a meta-heuristic technique for the feature selection. Further, this paper has been classified into five sections. Next section describes the meta-heuristic techniques which are followed by a section describing the hybridization of meta-heuristic techniques. Then implementation and result section analyzes the performance of different

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

hybrid meta-heuristic techniques including ACO-PSO-GA, ACO-PSO, ACO-GA, PSO-GA, and GA-DE. Then the conclusion section describes the overall conclusion from the analysis.

2. META-HEURISTIC TECHNIQUES

Meta-heuristic techniques mimic the behavior of any natural species or process to solve complex problems efficiently. A number of meta-heuristic techniques have been proposed like Ant Colony Optimization (ACO) which exhibits the foraging behavior of ant, Genetic Algorithm that shows the behavior of chromosomes generation and depletion. Few state of art meta-heuristic techniques has been described in this section.

2.1 Ant Colony Optimization

This technique mimics the foraging behavior of the ant which includes finding the shortest path to the food source. The ant follows the path which is having maximum pheromone value while the pheromone depends upon the number of ants moved through the path[7]. The complete process of ACO can be described using figure 1.

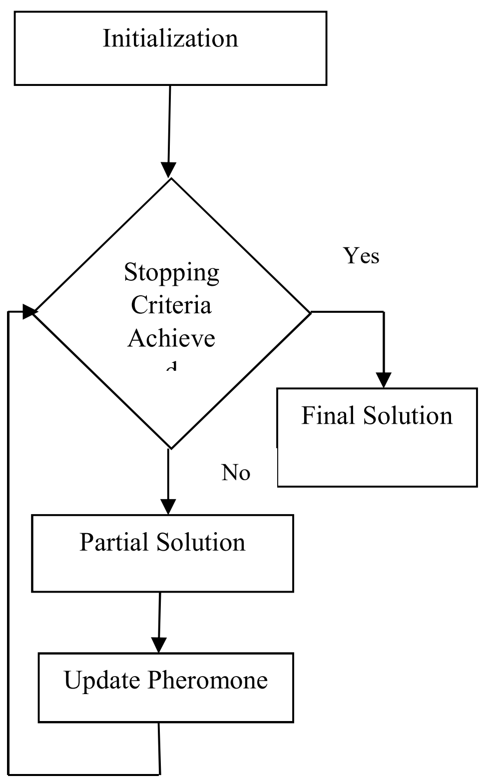

Figure 1: ACO Process

The complete process of ACO is described through figure 1 which provides the complete solution by updating the pheromone value.

2.2 Particle Swarm Optimization

This technique mimics the swarming and flocking behavior of the birds. This technique has a unique feature to update the solution as per the individual experience (local best) as well as the neighbor's experience (global best). This is the main reason for better convergence of technique towards global optima[8]. The overall process of PSO is described in figure 2.

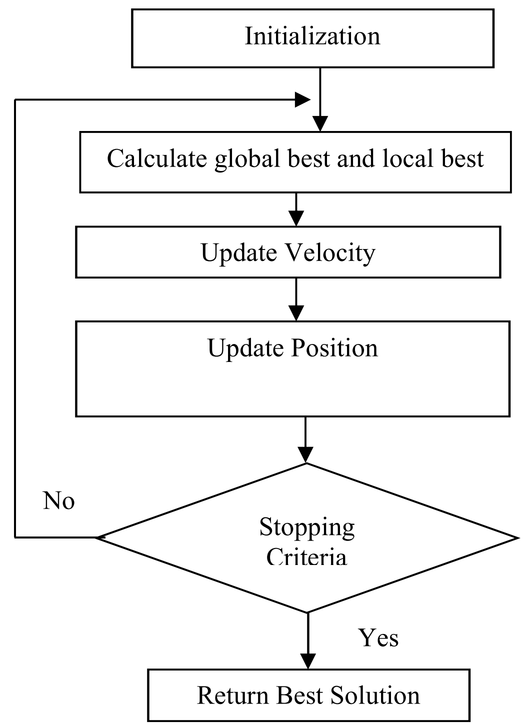

Figure 2: PSO Process

Figure 2 shows the complete process of PSO in which the particle updates its velocity and position based on local as well as global best solutions to achieve the global optima.

2.3 Genetic Algorithm

GA represents the natural occurring evolution process. This process represents, the generations of chromosomes that consist of genes which are reproduced as well as depleted through crossover, mutation operations[9]. The complete process is described in figure 3.

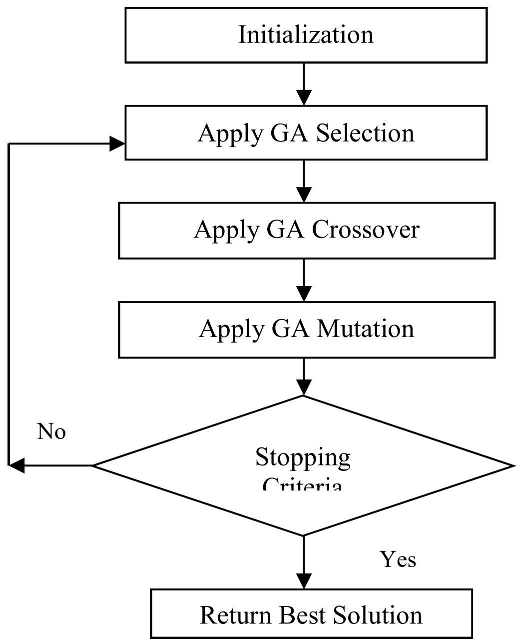

Figure 3: GA Process

The complete process of GA is described in figure 3 which shows all the operators including selection, crossover, and mutation applied to the population to generate the new population.

2.4 Differential Evolution

It is also a population-based technique like GA and uses the same operators i.e. selection, mutation and crossover.

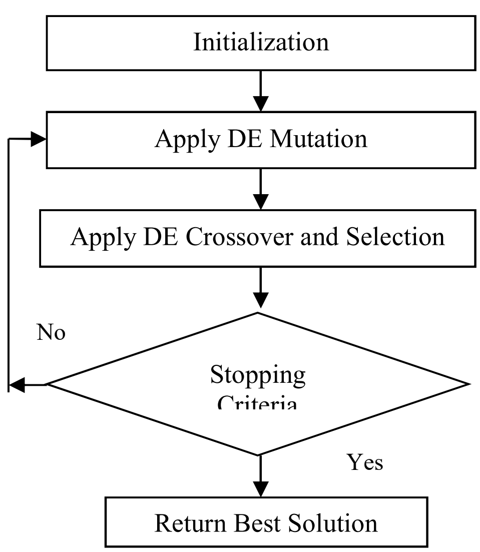

Figure 4: DE process

But the operation of GA especially mutation differs from the operations of DE[10][11]. It uses the concept of recombination to generate a faster solution. The complete process is shown in figure 4 which shows the flow of the process to generate the best solution.

Each meta-heuristic technique has its pros and cons. These techniques can be hybridized to improve individual performance which is discussed in the next section.

3. HYBRID META-HEURISTIC APPROACH

Different meta-heuristic approach converges towards the global optima but no approach guarantees the global optima. Several new meta-heuristic approaches like Dragonfly algorithm[12], binary gray wolf optimization[13], whale algorithm for optimization[2], ant ion optimization[14] are also proposed that performs better than the existing techniques due to better balancing between the exploration and exploitation phase. Moreover, several authors have hybridized various meta-heuristic techniques to improve the exploration and exploitation search by balancing the features of more than one meta-heuristic technique. This hybridization seems more promising. That's why this paper discusses and analyzes the different hybrid meta-heuristic techniques as follows.

3.1 Hybrid ACO-PSO Approach

The complete process of running ACO and PSO in parallel and selecting the best value by comparing the best solution of each is shown in figure 5. This algorithm performs better due to the merging of local search advantage of ACO and global perspective of PSO. Different authors have proposed various hybrid versions of ACO-PSO algorithms. This work focuses on one of the latest hybridizations of ACO-PSO which is proven to be effective by the author of [15], [16]. In this hybridization named as ACO-PSO1, the process of ACO and PSO is executed in parallel for each particle. However, the best position in each iteration is selected by comparing the best of ACO and PSO. The process is shown in figure 5.

The author of [15] gives another approach known as ACO-PSO2, which performs an initial search using the ACO. The initial search by ACO remains to continue until the fitness value gets stable. The

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

resultant of the ACO i.e. best subset found by the ACO act as the input to the PSO. Then, PSO updates the solution in each iteration and gives the best so shown in figure 6.

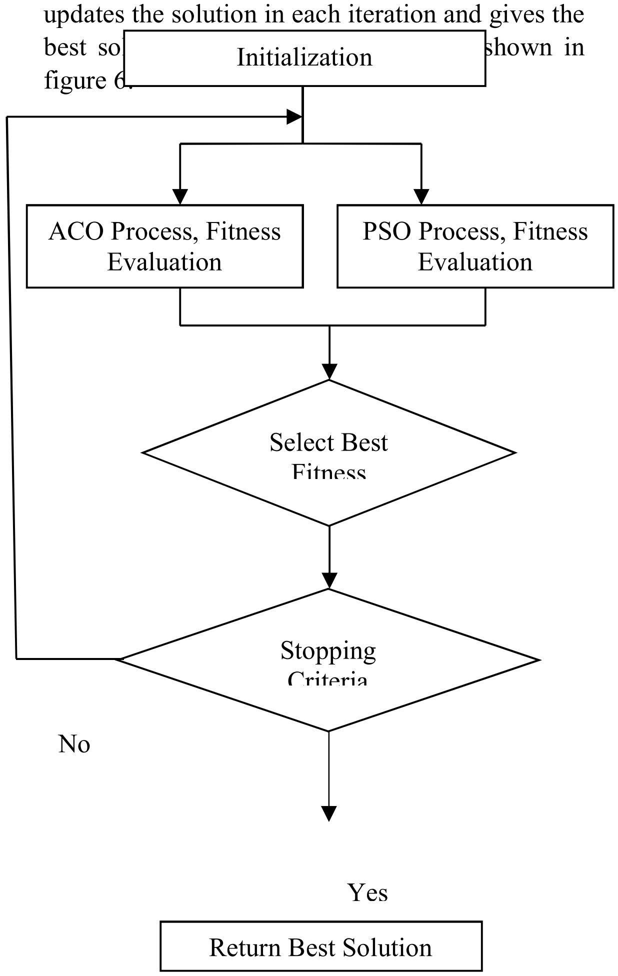

Figure 5: ACO-PSO1

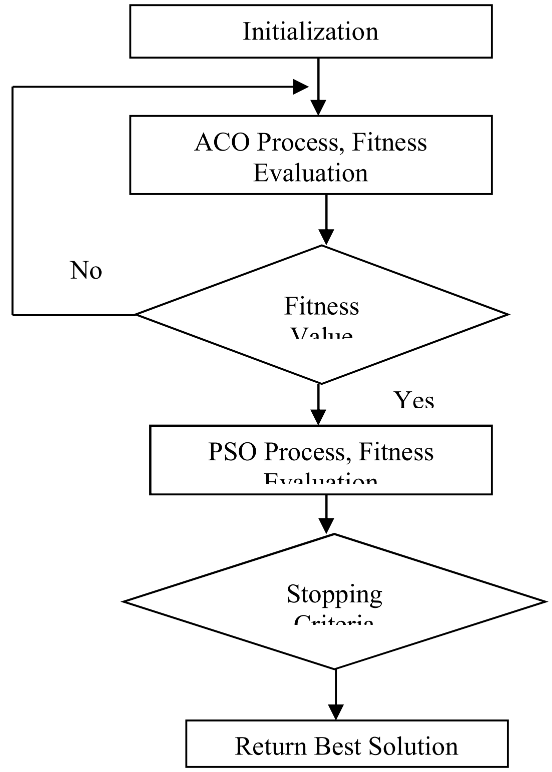

Figure 6: ACO-PSO2

Figure 6 presents the process of ACO-PSO2 which executes the ACO first to find the best solution which acts as an input to the PSO to find a better solution.

Another approach named as ACO-PSO3 given by the same author [15], hybridize the ACO and PSO by changing pheromone updating phenomena. The pheromone update is tied to the local and global search phenomena of PSO. In this approach, the pheromone update includes the best possible solution along with the local exploration as shown in figure 7.

Figure 7 shows the hybrid ACO-PSO3 process, in which the pheromone update considers the local as well as global results. The analysis of these hybrid algorithm has been done in the result section.

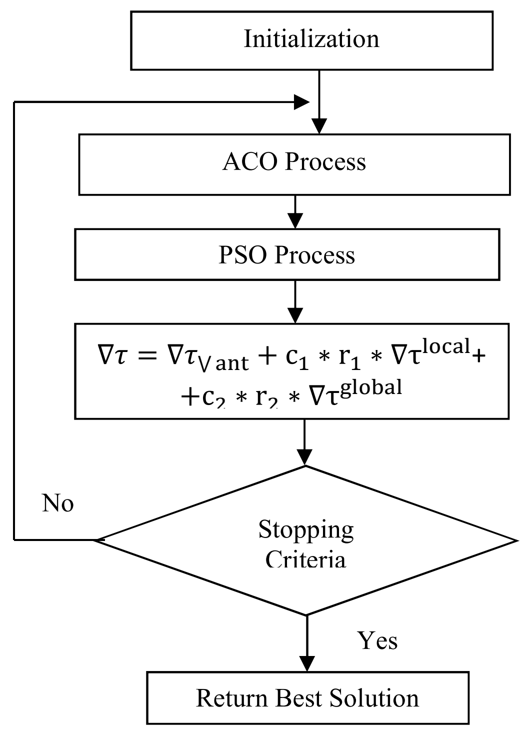

Figure 7: ACO-PSO3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3.2 Hybrid ACO-GA Algorithm

The author of [17]gives hybrid ACO-GA algorithm by running in parallel the ACO and the GA on the same population and evaluate both for the best results. The process is executed for the given minimum runs and updated the pheromone in each generation as shown in figure 8.

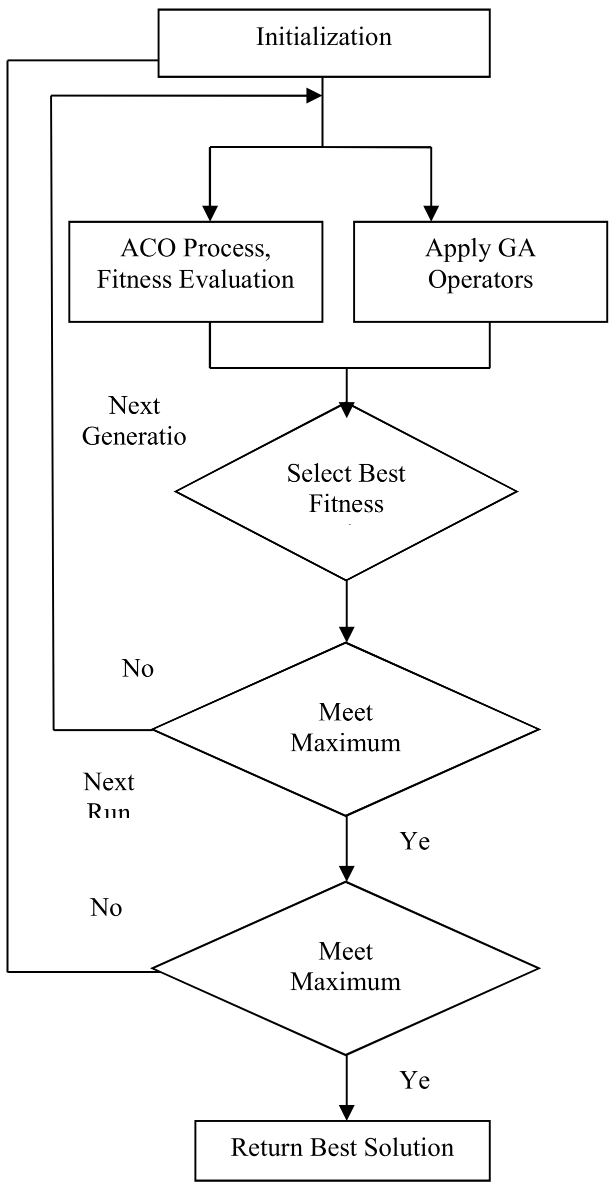

Figure 8: Hybrid ACO-GA

The figure 8 shows the hybrid process of ACO-GA algorithm, the algorithm iterates for given runs while in each iteration the algorithm executes for given generations. In each generation the process of ACO and GA operators is applied to find the solution. The better solution is saved and the pheromone value is updated for the next generation. This algorithm processes the better results as compared to the ACO as well as GA algorithm.

3.3 Hybrid PSO-GA Algorithm

The author of [18] hybrids the PSO and GA algorithm to avoid the local maxima problem of individual algorithms. The author has modified the weight updating mechanism to improve the global search. This algorithm starts with the random population and updates the position and velocity as per PSO algorithm to get the best value. The GA operator is applied to reproduce the population each step. It avoids premature convergence and improves the probability of convergence towards global optima. The complete process is described in figure 9.

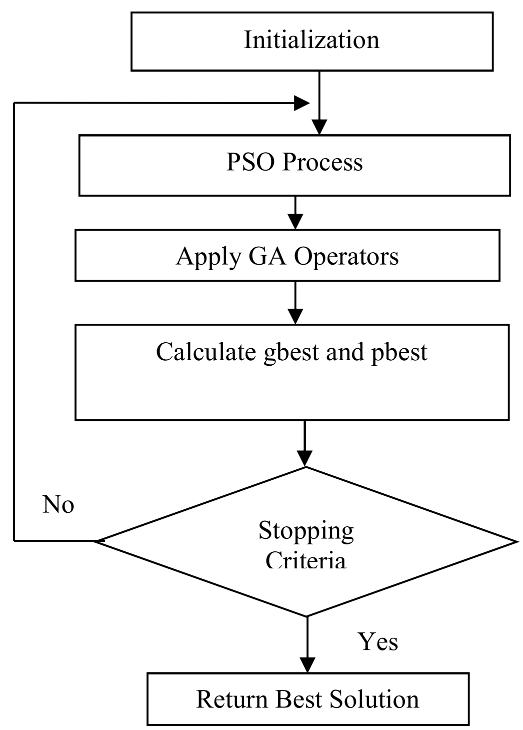

Figure 9: Hybrid PSO-GA algorithm

Figure 9 shows the hybrid PSO-GA algorithm in which the PSO process is applied to the random population to calculate the best values then the population is reproduced using the GA operators to recalculate the best value. The process gives better solution as compared to individual i.e. PSO and GA algorithms.

3.4 Hybrid GA-DE algorithm

GA-DE hybrid algorithm is given by the author of [19]. The author replaces the cross operation of GA with the differential vector perturbations to improve the balance of exploration and exploitation search. The complete process is shown in figure 10.

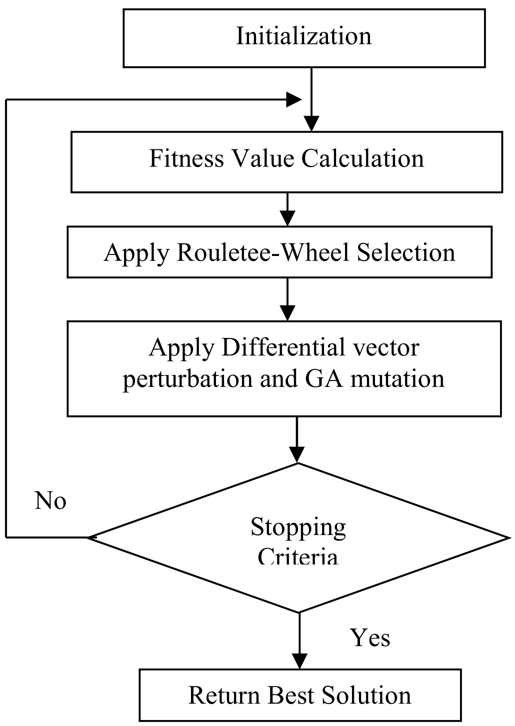

Figure 10: Hybrid GA-DE algorithm

Figure 10 denotes the process of hybrid GA-DE algorithm. The process clearly shows that the basic structure of the GA algorithm is adopted which includes the fitness function evaluation along with the roulette-wheel selection. Then the crossover step of GA has been replaced by the differential vector perturbation which is followed by the mutation step of GA. This process improves the performance by applying major perturbation in the crossover step and minor perturbation in the mutation step. The performance of the algorithm is better than the existing individual algorithms.

3.5 Hybrid ACO-PSO-GA Algorithm

In this, the author[20] hybridize all the three algorithms i.e. ACO, PSO and GA to remove their corresponding weakness. The problem with GA is the random initiation while the PSO mainly focuses on exploitative search. The ACO algorithm performs better explorative search. So, all the three combined to converge towards the global optima. The process of hybrid ACO-PSO-GA is elaborated in figure 11.

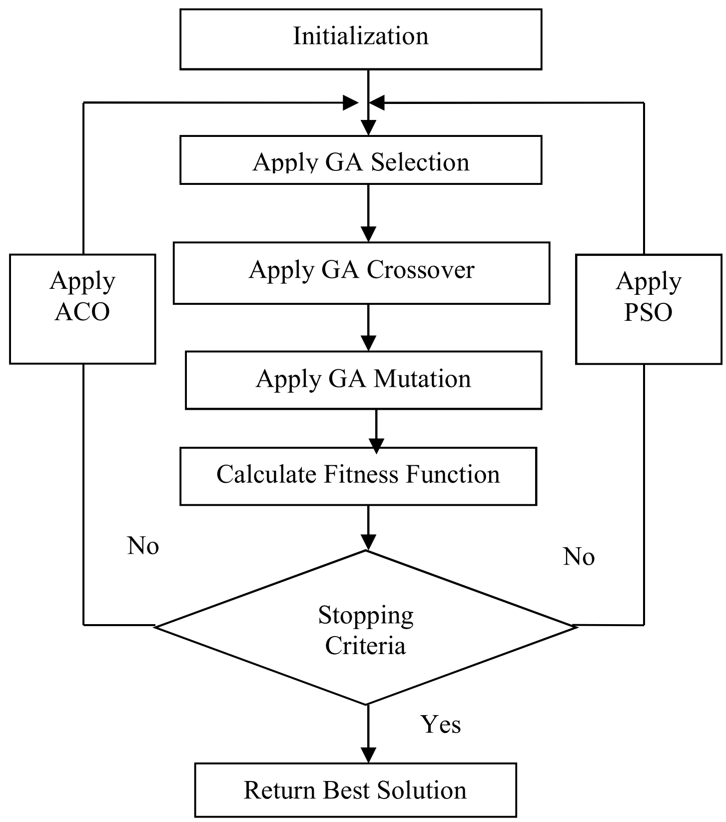

Figure 11: Hybrid ACO-PSO-GA Algorithm

The figure 11 represents the steps of hybrid ACO-PSO-GA algorithm. In this algorithm the random population is given as input to the GA then the ACO and PSO process reproduces the population which is updated by the GA operators. Here, ACO with GA improves the exploratory search and remove the repeated path while the ACO with PSO improves the exploitation search. The process is repeated until the stopping criteria achieved. This process gives better performance as compared to the ACO, PSO and the GA algorithm.

4. RESULT AND DISCUSSION

This work analyzes the hybrid techniques on 7 different datasets. The detail of datasets including dataset name along with its resource, attributes and instances are given in table 1. Moreover, the total

Journal of Theoretical and Applied Information Technology
31st August 2019. Vol.97. No 16

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

number of instances available in the dataset is also given in table 1.

Table 1: Datasets Description

S. No.	Dataset Name	Attributes	Instance	Total Elements	Resource
1	AUSTRALIAN	14	690	9660	UCI Repository
2	WDBC	30	569	17070	
3	PIMA	8	768	6144	
4	ARRHYTHMIA	452	279	126108	
5	ADVERTISEMENT	1558	3279	5108682	
6	HAR	561	10299	5777739	
7	MADELON	500	2000	1000000	

Table 1 shows the description of the datasets used in the work to analyze the performance of hybrid meta-heuristic techniques for classification. The SVM classifier is used for classification purposes. Here, the 10 fold cross-validation is used for the training and testing target. The analysis has been done using the accuracy, sensitivity, specificity and selected feature ratio described below:

i) Accuracy

Accuracy represents the number of instances correctly classified by the classifier. It is given as:

$$A = \frac{1}{n} * Correctly\_classified\_intsnaces \quad (1)$$

Here, n is the number of instances in the dataset

ii) Sensitivity

It gives the correctly classified true instances. In other words, it is sensitivity to the correct classification. The sensitivity can be represented by (2).

$$Senstivity = \frac{TP}{TP+FN} \quad (2)$$

Here TP, FN shows true positive and false negative respectively.

iii) Specificity

It is the correctly classified negative instances. It can be given by (3).

$$Specificity = \frac{TN}{TN+FP} \quad (3)$$

Here TN, FP denotes true negative and false positive respectively.

iv) Selected feature ratio

It is the ratio of the number of features selected to the total number of features. It is given by (4):

$$SFR = \frac{NumberofselectedFeatures}{n} \quad (4)$$

Here, n is the number of features in the dataset.

4.1 Performance analysis

The analysis to compare the performance of hybrid meta-heuristic algorithms i.e. ACO-PSO, ACO-GA, PSO-GA, GA-DE and ACO-PSO-GA using the parameters described above on 6 datasets given in table 1 has been done in this section. The results are evaluated by executing the algorithms 20 times and taking the average of results. The initial parameter setting has been done as per the referred paper for each algorithm. The SVM classifier is used with RBF kernel having sigma=15. The accuracy is given in table 2.

Table 2: Comparison of Accuracy on Hybrid Meta-heuristic techniques

Dataset	ACO-PSO	GA-DE	PSO-GA	ACO-GA	ACO-PSO-GA
AUSTRALIAN	0.852	0.852	0.849	0.849	0.863
WDBC	0.968	0.979	0.920	0.990	1.003
PIMA	0.788	0.793	0.775	0.791	0.817
ARRHYTHMIA	0.537	0.588	0.559	0.586	0.669
ADVERTISEMENT	0.929	0.965	0.949	0.955	0.986
MADELON	0.567	0.585	0.596	0.596	0.616
HAR	0.894	0.910	0.849	0.939	0.971

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The comparison of accuracy for hybrid meta-heuristic technique can also be shown graphically.

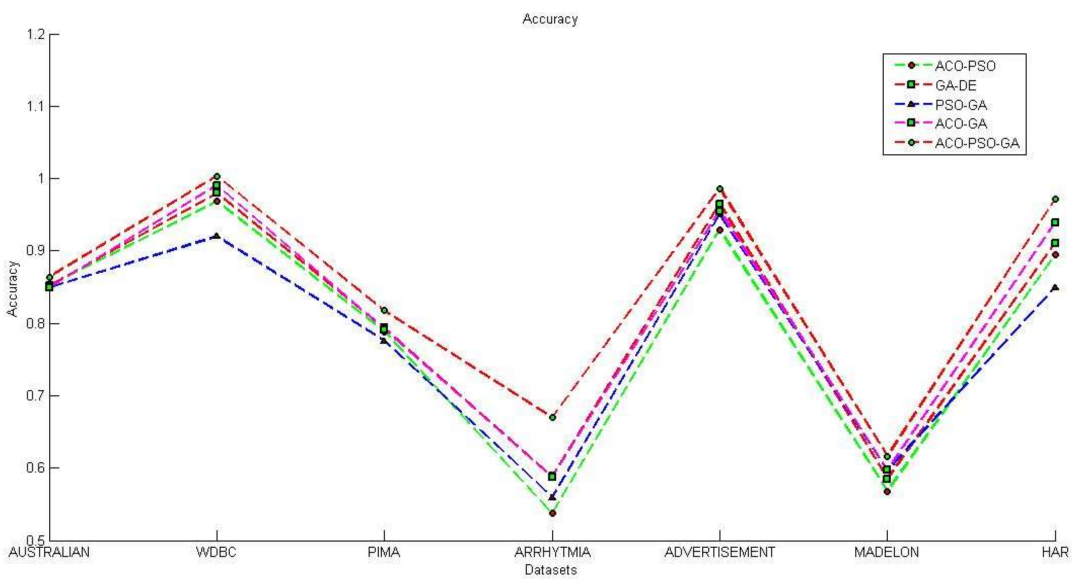

Figure 12: Analysis of Accuracy on Different datasets

Figure 12 compares the classification accuracy on different datasets achieved by the hybrid algorithms. The hybrid ACO-PSO-GA algorithm exhibits more accuracy on each dataset as compared to other algorithms. This is due to balancing of exploration phase (due to ACO) and exploitation phase (PSO) by the GA. The comparison of ratio for the number of features selected to the total number of features is given in table3.

Table 3: Comparison of Selection Size Ratio on Hybrid Meta-heuristic techniques

Dataset	ACO-PSO	GA-DE	PSO-GA	ACO-GA	ACO-PSO-GA
AUSTRALIAN	0.868	0.468	0.296	0.368	0.654
WDBC	0.744	0.387	0.177	0.277	0.444
PIMA	0.911	0.611	0.386	0.261	0.636
ARRHYTHMIA	0.318	0.259	0.185	0.429	0.124
ADVERTISEMENT	0.639	0.644	0.255	0.725	0.543
MADELON	0.5162	0.4414	0.4752	0.5216	0.3112
HAR	0.5108	0.457	0.2785	0.5124	0.4011

The comparison shown in table 3 is also represented graphically in figure 13.

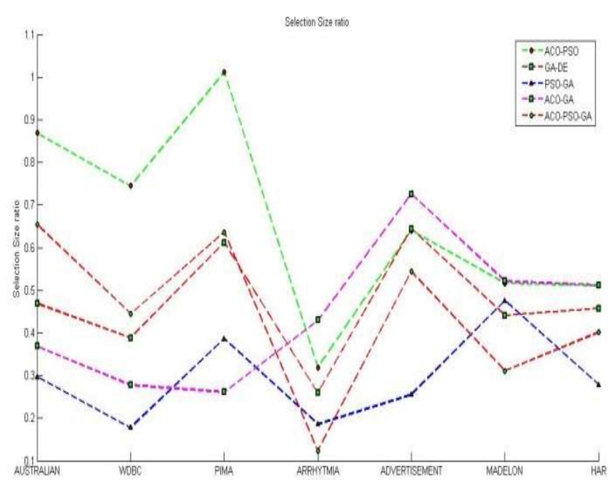

Figure 13: Analysis of Selection size ratio on Different datasets

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 13 compares the selection size ratio of the hybrid algorithm on different datasets. It can be seen that the selection size of the ACO-PSO-GA is high as compared to other few hybrid algorithms. It means higher accuracy of ACO-PSO-GA algorithm costs the number of features.

Table 4: Comparison of Sensitivity on Hybrid Meta-heuristic techniques

Dataset	ACO-PSO	GA-DE	PSO-GA	ACO-GA	ACO-PSO-GA
AUSTRALIAN	0.7486	0.7525	0.7486	0.7486	0.7681
WDBC	0.9162	0.9416	0.9707	0.9707	0.9889
PIMA	0.765	0.7795	0.7936	0.840	0.826
ARRHYTHMIA	0.777	0.837	0.822	0.822	0.944
ADVERTISEMENT	0.921	0.933	0.940	0.965	0.979
MADELON	0.567	0.575	0.530	0.597	0.587
HAR	0.602	0.973	0.976	0.964	0.983

The comparison of sensitivity on various datasets is done in table 4 is also shown graphically in figure 14.

The comparison clearly shows the sensitivity of the hybrid ACO-PSO-GA algorithm is better as compared to other hybrid techniques. However, the exception is shown for the PIMA dataset.

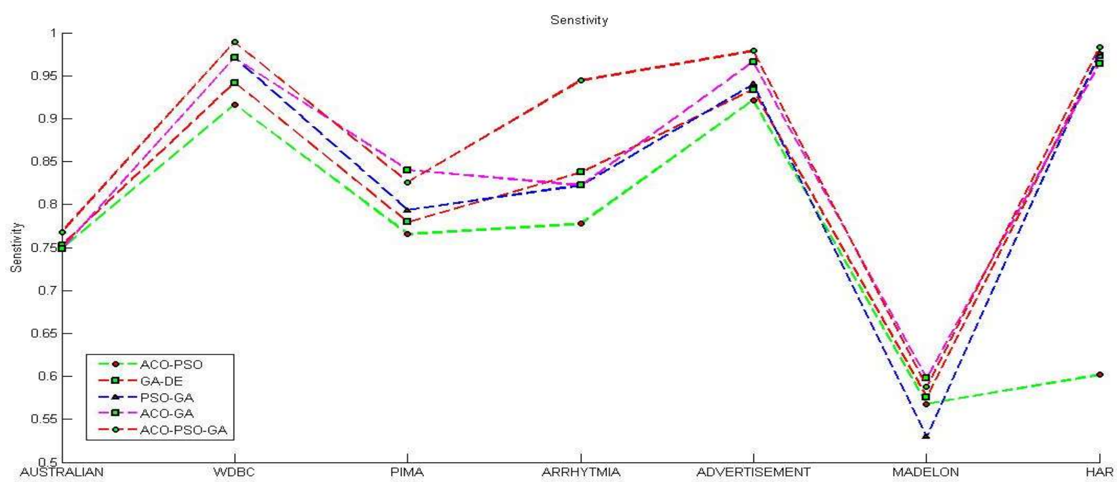

Figure 14: Analysis of Sensitivity on Different datasets

The comparison of the specificity on various datasets is shown in table 5.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 5: Comparison of Specificity on Hybrid Meta-heuristic techniques

Dataset	ACO-PSO	GA-DE	PSO-GA	ACO-GA	ACO-PSO-GA
AUSTRALIAN	0.931	0.925	0.923	0.923	0.931
WDBC	0.954	0.961	0.87	0.965	0.977
PIMA	0.763	0.746	0.65	0.6018	0.720
ARRHYTHMIA	0.340	0.431	0.450	0.417	0.527
ADVERTISEMENT	0.799	0.842	0.842	0.738	0.738
MADELON	0.521	0.548	0.575	0.549	0.597
HAR	0.575	0.968	0.303	0.950	0.976

The comparison is also done graphically shown in figure 15.

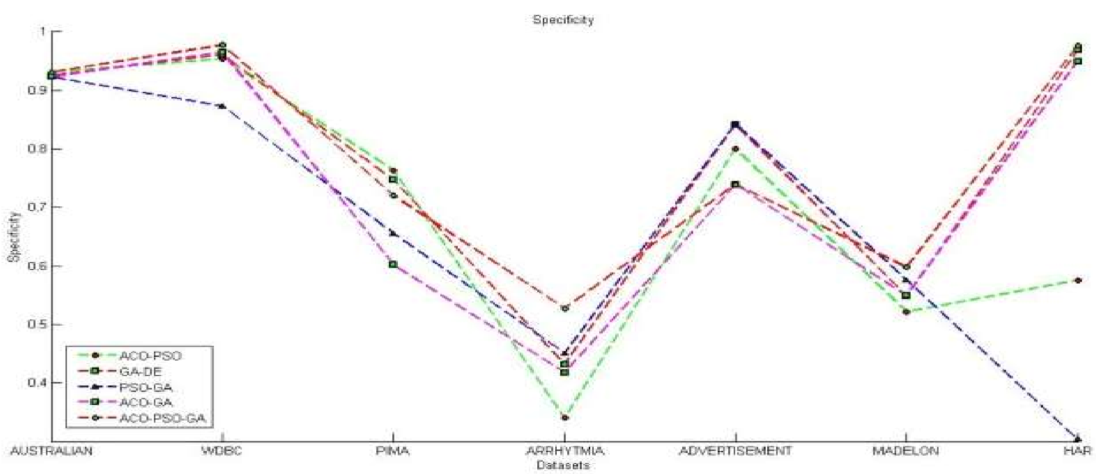

Figure 15: Analysis of Specificity on Different datasets

The comparison in figure 15 clearly denotes that the specificity of hybrid ACO-PSO_GA algorithm is better as compared to other hybrid algorithms due to a more balanced exploration and exploitation search.

5. CONCLUSION

This paper analyzes different hybridization of ACO, PSO, GA and DE for the feature selection to improve the classification performance. The analysis has been done on seven different datasets with varying size using accuracy, sensitivity, and specificity and selection size ratio as the parameters. The analysis clearly shows that the hybrid ACO-PSO-GA performs better as compared to other hybrid algorithms due to the combined advantages of individual algorithms. However, the selection size ratio of the algorithm is higher as compared to other algorithms. In the future, the hybridization can be extended to improve the performance of the algorithm.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

REFERENCES

[1] P. Lopez-Garcia, A. D. Masegosa, E. Osaba, E. Onieva, and A. Perallos, "Ensemble classification for imbalanced data based on feature space partitioning and hybrid metaheuristics," *Applied Intelligence*, vol. 2019, pp. 1–16, 2019.

[2] M. M. Mafarja and S. Mirjalili, "Hybrid Whale Optimization Algorithm with simulated annealing for feature selection," *Neurocomputing*, vol. 260, pp. 302–312, 2017.

[3] G. I. Sayed, A. E. Hassanien, and A. T. Azar, "Feature selection via a novel chaotic crow search algorithm," *Neural Computing and Applications*, vol. 31, pp. 171–188, 2019.

[4] A. M. Anter, A. T. Azar, and K. M. Fouad, "Intelligent Hybrid Approach for Feature Selection," *Springer Nature*, vol. 3, pp. 71–79, 2019.

[5] M. N. Ab Wahab, S. Nefti-Meziani, and A. Atyabi, "A comprehensive review of swarm optimization algorithms," *PLoS ONE*, vol. 10, no. 5, pp. 1–36, 2015.

[6] I. Boussaïd, J. Lepagnot, and P. Siarry, "A survey on optimization metaheuristics," *Information Sciences*, vol. 237, no. February, pp. 82–117, 2013.

[7] M. Guntsch and M. Middendorf, "A Population Based Approach for ACO," pp. 72–81, 2007.

[8] "Empirical study of PSO 1999.pdf." .

[9] A. Dastanpour and R. A. R. Mahmood, "Feature selection based on genetic algorithm and SupportVector machine for intrusion detection system," *International Conference on Informatics Engineering & Information Science*, no. September 2014, pp. 169–181, 2013.

[10] R. Storn and K. Price, "Differential Evolution - A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, 1997.

[11] Z. Huang and Y. Chen, "An Improved Differential Evolution Algorithm Based on Adaptive Parameter," *Journal of Control Science and Engineering*, vol. 2013, pp. 1–5, 2013.

[12] M. Mafarja *et al.*, "Binary dragonfly optimization for feature selection using time-varying transfer functions," *Knowledge-Based Systems*, vol. 161, pp. 185–204, 2018.

[13] E. Emary, H. M. Zawbaa, and A. E. Hassanien, "Binary grey wolf optimization approaches for feature selection," *Neurocomputing*, vol. 172, no. 2016, pp. 371–381, 2016.

[14] M. M. Mafarja and S. Mirjalili, "Hybrid binary ant lion optimizer with rough set and approximate entropy reducts for feature selection," *Soft Computing*, vol. 2018, pp. 1–17, 2018.

[15] K. Menghour and L. Souici-Meslati, "Hybrid ACO-PSO based approaches for feature selection," *International Journal of Intelligent Engineering and Systems*, vol. 9, no. 3, pp. 65–79, 2016.

[16] N. P. Holden and A. A. Freitas, "A hybrid PSO/ACO algorithm for classification," p. 2745, 2007.

[17] S. Nemati, M. E. Basiri, N. Ghasem-Aghaee, and M. H. Aghdam, "A novel ACO-GA hybrid algorithm for feature selection in protein function prediction," *Expert Systems with Applications*, vol. 36, no. 10, pp. 12086–12094, 2009.

[18] S. M.Uma, K. Rajiv Gandhi, E. Kirubakaran, and D. E. K. Dr.E.Kirubakaran, "A Hybrid PSO with Dynamic Inertia Weight and GA Approach for Discovering Classification Rule in Data Mining," *International Journal of Computer Applications*, vol. 40, no. 17, pp. 32–37, 2012.

[19] W. Y. Lin, "A GA-DE hybrid evolutionary algorithm for path synthesis of four-bar linkage," *Mechanism and Machine Theory*, vol. 45, no. 8, pp. 1096–1107, 2010.

[20] J. H. Tam, Z. C. Ong, Z. Ismail, B. C. Ang, and S. Y. Khoo, "A new hybrid GA−ACO−PSO algorithm for solving various engineering design problems," *International Journal of Computer Mathematics*, vol. 0, no. 0, pp. 1–37, 2018.