Chapter 7

Adaptive Genetic Algorithm with Modified Operator

7.1 Adaptive Genetic Algorithm with Modified Operator: AGA-MO

In this work, genetic algorithm (GA) based novel method, Adaptive Genetic Algorithm with Modified Operator(AGA-MO), has been developed and has been implemented for simultaneous feature selection. The AGA-MO method feature selection method uses a population of reducts. For each of the candidate reduct (chromosome), fitness function is evaluated as a function of (i) fuzzy rough dependency measure, $\gamma_P^{'}(Q)$, and (ii) cardinality, $|R|$, of the reduct R (Equation 4.5). Fuzzy rough dependency measure, $\gamma_P^{'}(Q)$, for reduct R corresponding to class label, Q, is calculated using RST and L-FRFS measures. The proposed method of AGA-MO finds optimal reduct (chromosome) using these fitness function values. In this work, a reduct is represented as a binary string, which shall be called as a chromosome for implementing the feature selection using AGA-MO. In this string, features corresponding to the indices of each '1' are selected features and those corresponding to the indices of each '0' are not selected in a chromosome. Thus, the length of the string (chromosome) remains same as the number of features in the dataset being used irrespective of the number of features selected for any reduct.

Algorithm 11 AGA-MO

1: **procedure** AGA-MO()
2:     **procedure** INITIALIZATION()
3:         Initialize N, F, CR and d
4:         **for** $j = 1$ to N **do**
5:             **for** $j = 1$ to d **do**
6:                 for DS-initialization use *Algorithm 4*, from Chapter 4.
7:                 $f_i \leftarrow$ fitness value at $\bar{x}_i$
8:             **end for**
9:         **end for**
10:     **end procedure**
11:     **while** termination condition is not satisfied **do**
12:         **procedure** PARENT SELECTION()
13:             **for** $i = 1$ to N **do**
14:                 $P_1(i) \leftarrow \bar{x}_i$
15:                 $P_2(i) \leftarrow$ randomly select chromosome from among the population
16:                 **while** $P_1(i) == P_2(i)$ **do**
17:                     $P_2(i) \leftarrow$ randomly select chromosome from among the population
18:                 **end while**
19:             **end for**
20:         **end procedure**
21:         **procedure** XOR-MUTATION(P_2)
22:             **for** $i = 1$ to N **do**
23:                 // calculate total number of mutation points //
24:                 $r_1 \leftarrow$ randomly select integer from 1 to N
25:                 $r_2 \leftarrow$ randomly select integer from 1 to N
26:                 $l_i \leftarrow \text{floor}(F \times \text{Sum}\{XOR(\bar{x}_{r_1}, \bar{x}_{r_2})\})$
27:                 randomly select l_i mutation point in $P_2(i)$
28:                 Change the bit value of every mutation point
29:             **end for**
30:         **end procedure**
31:         **procedure** RECOMBINATION(P_1, P_2)
32:             // Binomial Crossover //
33:             $P_3 \leftarrow \text{Binxov}\{P_1, P_2, CR\}$
34:         **end procedure**
35:         **procedure** SELECTION(P_1, P_3)
36:             **for** $i = 1$ to N **do**
37:                 **if** $f(P_1(i)) > f(P_3(i))$ **then**
38:                     // P_1 is chosen as a prominent parent because sharing of P_1 in generation of offspring is more than P_2. //
39:                     $\bar{x}_i \leftarrow P_1(i)$
40:                 **else**
41:                     $\bar{x}_i \leftarrow P_3(i)$
42:                 **end if**
43:             **end for**
44:         **end procedure**
45:     **end while**
46:     **return** Best Solution
47: **end procedure**

7.1.1 Algorithm

A population of size N has been initialized in which dimension of each of the chromosomes is d (statements: 2 to 9) and fitness values for the chromosomes are also calculated (statements: 7). For i^{th} chromosome, two parents, $P_1(i)$ and $P_2(i)$, are to be chosen. $P_1(i)$ is $\bar{x}_i$ itself and $P_2(i)$ is randomly selected from among the current population of N chromosomes (statements: 11-19) ensuring that $P_2(i) \neq P_1(i)$. (statements 16-18).

Two chromosomes $\bar{x}_{r_1}$ and $\bar{x}_{r_2}$ are chosen randomly from the current population (statement: 24-25). For i^{th} chromosomes statement # 42 returns an integer l_i proportional to the closeness of two randomly chosen chromosomes, $\bar{x}_{r_1}$ and $\bar{x}_{r_2}$. Now in $P_2(i)$, l_i bits are randomly mutated (statements: 27-28).

For each i, a binomial crossover between corresponding $P_1(i)$ and $P_2(i)$ is performed in bit-by-bit fashion using predefined Crossover Rate, CR, to get new chromosome $P_3(i)$. After crossover is done, we get two populations, P_1 and P_3. Now, to get updated i^{th} chromosome in the next population, $\bar{x}_i^{k+1}$, that chromosome is chosen from the pair $P_1(i)$ and $P_3(i)$), which has higher fitness value.

This algorithm AGA-MO is modification of GA in such a way that the mutation operator is adaptively dependent on the distribution of particles in a population. When particles are diverse, the probability of mutation is high and when the population is converging, the mutation is reduced.

7.2 Experiments, Results and Discussions

The performance of the proposed AGA-MO method, has been compared with previously introduced BO method. eGA method has also been implemented for feature selection because AGA-MO is supposed as an enhanced version of GA. Initialization of population in case of eGA and AGA-MO has been implemented using DS-initialization. All the benchmark dataset are taken from the UCI data repository of machine learning, [77]. Table 4.2 tabulates number of objects and features for each of the dataset used in the present work.

7.2.1 AGA-MO method with RST measure

The results of the feature selection are presented in Table 7.1 in the terms of number of features in the reducts suggested by each of the methods. For each of the algorithm, one run is of 100 generation with population size of 100. In the proposed AGA-MO algorithm, the parameters F and CR are taken as 0.7 and 0.9 respectively. For Elitist Genetic Algorithm (eGA), elitism percentage, crossover rate and mutation rate are chosen as 20%, 0.50-0.70 and 0.005 respectively.

For the 12 runs of each method for any dataset, best feature subset (reduct) size, mean feature subset size (reduct) and standard deviation (s.d.) among the reduct sizes have been presented. To demonstrate that the proposed feature selection method applying AGA-MO is successful in identifying the noisy (irrelevant) and redundant (superfluous) features, classification accuracies have also been evaluated and discussed. To validate the performance of AGA-MO, the results of feature selection using this algorithm has been compared with results of feature selection using previously introduced Butterfly Optimizer (BO) and Elitist Genetic Algorithm (eGA) [16]. To validate that the accuracy of the reducts obtained from the proposed method, AGA-MO, are acceptable, Table 7.2 presents classification accuracies using three classifiers viz. J48 [59], JRip [79] and PART [80].

From Table 7.1 it is seen that in the case of *Cleveland*, *Ecoli*, *Glass*, *Soybean-small* and *Wine* dataset, AGA-MO is as good as other methods. In the case of *Ionosphere* and *Lung*, t-test and Wilcoxon test suggest that AGA-MO performs better than eGA in terms of reduct size. In the case of *LSVT* t-test and Wilcoxon test suggest that AGA-MO performs better than all other methods (BO and eGA) in terms of reduct size. In the case of AGA-MO, Table 7.1 shows that reduct size is minimum for all the dataset used. Further, in Table 7.2 t-test and Wilcoxon test results show that in case of AGA-MO, classification accuracy is comparable to BO and eGA for all the dataset. In the case of AGA-MO maximum accuracy is more than that of other methods for all the dataset, which means that AGA-MO is able to identify and remove the noisy and irrelevant features which otherwise lower the classification accuracy.

Tables 7.1 and 7.2 show best results reported in literature in terms of Mean Subset size (denoted as MSS) and classification accuracy (denoted as CA). Further, it is evident from the Tables 1.1, 7.1 and 7.2 that the performance of AGA-MO is better than that of

Table 7.1: AGA-MO with RST measure: Comparison of reduct size with statistical t-test, and Wilcoxon test. [Statistical t-test is denoted as 'S'. Wilcoxon test is denoted as 'WT'. Mean Subset Size is denoted as 'MSS'.]

Dataset (Total features)	Feature Selection Method	Feature Subset size				Best MSS reported in literature
		Min	Mean(s.d)	S	W.T.	
Cleveland(13)	eGA	3	3(0)	-	-	7.81 [30]
	BO	3	3(0)	-	-	
	AGA-MO	**3**	3(0)			
Ecoli(7)	eGA	3	3(0)	-	-	3 [52]
	BO	3	3(0)	-	-	
	AGA-MO	**3**	3(0)			
Glass(9)	eGA	2	2(0)	-	-	8.44 [30]
	BO	2	2(0)	-	-	
	AGA-MO	**2**	2(0)			
Ionosphere(34)	eGA	2	2.16(0.4)	v	v	7.3 [30]
	BO	2	2(0)	-	-	
	AGA-MO	**2**	2(0)			
Lung(56)	eGA	3	3.41(0.5)	v	v	NA
	BO	3	3(0)	-	-	
	AGA-MO	**3**	3(0)			
Soybean small(35)	eGA	2	2(0)	-	-	2 [29]
	BO	2	2(0)	-	-	
	AGA-MO	**2**	2(0)			
Wine(13)	eGA	2	2(0)	-	-	2 [52]
	BO	2	2(0)	-	-	
	AGA-MO	**2**	2(0)			
LSVT(310)	eGA	7	11.16(2.5)	v	v	NA
	BO	4	8.34(2.2)	v	v	
	AGA-MO	**1**	**1.5(1)**			

state-of-the-art methods suggested in literature. While using rough dependency measure as fitness function, these method provides smaller reduct with comparable or more accuracy than the existing best method reported in literature, for all the dataset. AGA-MO produces the stable result with zero or very low standard deviation.

Table 7.3 shows the results for Friedman test giving the ranking of the performance of AGA-MO, BO and eGA. It is observed that AGA-MO ranks the best as compared to BO and eGA.

Table 7.2: AGA-MO with RST measure: Comparison of classification accuracy with statistical t-test, and Wilcoxon test. [Classification accuracy is denoted as 'CA'. Statistical t-test is denoted as 'S'. Wilcoxon test is denoted as 'WT'.]

| Dataset (Total features) | Feature Selection Method | Classification Accuracy (CA) | | | | | | | | | | | | Best CA reported in literature |
| | | Classifier: J48 | | | | Classifier: JRip | | | | Classifier: PART | | | | |
		Max	Mean(s.d)	S	W.T.	Max	Mean(s.d)	S	W.T.	Max	Mean(s.d)	S	W.T.	
Cleveland(13)	eGA	55.44	53.02(2.32)	-	-	56.1	53.65(1.18)	-	-	54.78	51.78(2.89)	-	-	52.6 [30]
	BO	55.44	52.67(1.62)	-	-	**56.1**	52.79(1.14)	-	-	54.78	51.56(2.21)	-	-	
	AGA-MO	55.44	53.02(2.1)			56.1	53.46(1.3)			54.78	51.83(2.41)			
Ecoli(7)	eGA	79.46	76.41(4.16)	-	-	80.95	76.06(5.84)	-	-	80.95	76.11(6.1)	-	-	77.38 [52]
	BO	79.46	77.22(2.51)	-	-	80.95	77.59(2.27)	-	-	80.95	77.31(1.1)	-	-	
	AGA-MO	79.46	78.66(1.04)			**80.95**	79.11(2.56)			80.95	79.38(2.05)			
Glass(9)	eGA	66.36	61.63(6.03)	-	-	63.08	57.78(5)	-	-	68.22	61.48(7.13)	-	-	65.14 [30]
	BO	66.36	62.46(2.02)	-	-	64.95	59.17(3.01)	-	-	68.22	61.64(3.17)	-	-	
	AGA-MO	66.36	63.27(2.74)			64.95	60.54(1.87)			**68.22**	62.8(4.03)			
Ionosphere(34)	eGA	88.89	86.01(2.1)	-	-	88.89	86.24(2.17)	-	-	87.75	85.17(2.09)	-	-	86.17 [30]
	BO	88.89	86.71(2.31)	-	-	88.89	86.53(1.63)	-	-	87.75	85.85(2.9)	-	-	
	AGA-MO	**88.89**	87.41(1.27)			88.89	87.6(1.71)			87.75	86.74(1.66)			
Lung(56)	eGA	87.5	85.67(2.81)	-	-	87.5	85.41(3.6)	-	-	87.5	83.59(4.24)	-	-	NA
	BO	87.5	86.27(2.11)	-	-	87.5	85.82(2.16)	-	-	87.5	83.9(2.32)	-	-	
	AGA-MO	**87.5**	86.45(1.54)			87.5	86.71(1.41)			87.5	84.37(2.98)			
Soybean small(35)	eGA	100	100(0)	-	-	100	100(0)	-	-	100	100(0)	-	-	100 [29]
	BO	100	100(0)	-	-	100	100(0)	-	-	100	100(0)	-	-	
	AGA-MO	**100**	100(0)			100	100(0)			100	100(0)			
Wine(13)	eGA	94.94	78.92(11)	-	-	90.45	76.91(9.32)	-	-	93.82	78.55(10.6)	-	-	90.44 [52]
	BO	94.94	85.39(1.31)	-	-	90.45	86.37(2.45)	-	-	93.82	85.45(1.73)	-	-	
	AGA-MO	**94.94**	90.11(3.48)			90.45	87.02(3.44)			93.82	89.55(3.43)			
LSVT(310)	eGA	78.57	74.2(2.47)	-	-	80.95	75.52(3.97)	-	-	76.98	72.74(2.75)	-	-	NA
	BO	80.16	77.12(2.12)	-	-	80.95	77.85(1.87)	-	-	76.98	76.54(2.35)	-	-	
	AGA-MO	80.16	78.1(1.8)			**80.95**	78.43(2.34)			80.16	78.3(1.7)			

Table 7.3: Friedman ranking with RST measure

Methods	FR	Rank
AGA-MO	1.3281	1
BO	2.0625	2
eGA	2.6094	3

7.2.2 AGA-MO method with L-FRFS measure

The results of the feature selection are presented in Table 7.4 in terms of the number of features in the reducts suggested by each of the methods. For each of the algorithm, one run is of 100 generation with population size of 100. In the proposed AGA-MO algorithm, the parameters F and CR are taken as 0.7 and 0.9 respectively. For Elitist Genetic Algorithm (eGA), elitism percentage, crossover rate and mutation rate are chosen as 20%, 0.50-0.70 and 0.005 respectively.

For the 12 runs of each method for any dataset, best feature subset (reduct) size, mean feature subset size (reduct) and standard deviation (s.d.) among the reduct sizes have been presented.To demonstrate that the proposed feature selection method applying AGA-MO is successful in identifying the noisy (irrelevant) and redundant (superfluous) features, classification accuracies have also been evaluated and discussed. To validate the performance of AGA-MO, the results of feature selection using this algorithm has been compared with results of feature selection using previously introduced Butterfly Optimizer (BO) and Elitist Genetic Algorithm (eGA) [16]. To validate that the accuracy of the reducts obtained from the proposed method, AGA-MO, are acceptable, Table 7.5 presents classification accuracies using three classifiers viz. J48 [59], JRip [79] and PART [80].

From Table 7.4 it is seen that in the case of *Cleveland*, *Ecoli*, *Glass*, *Soybean-small* and *Wine* dataset, AGA-MO is as good as other methods. In the case of *Ionosphere*, *Lung* and *LSVT* t-test and Wilcoxon test suggest that AGA-MO performs better than eGA and BO in terms of reduct size. In the case of AGA-MO, Table 7.4 shows that reduct size is minimum for all the dataset used. Further, in Table 7.5 t-test and Wilcoxon test results show that in case of AGA-MO, classification accuracy is comparable to BO and eGA for *Cleveland*, *Ecoli*, *Glass* and *Soybean small* the dataset. Table 7.5 shows that

for *Ionosphere*, *Lung*, *Wine* and *LSVT* t-test and Wilcoxon test suggest that AGA-MO performs better than eGA in terms of classification accuracy. It is observed from Table 7.5 that for *Ionosphere*, *Lung*, *Wine* and *LSVT* t-test and Wilcoxon test suggest that AGA-MO is as good as BO in terms of classification accuracy.

In the case of AGA-MO maximum accuracy is more than that of other methods for all the dataset, which means that AGA-MO is able to identify and remove the noisy and irrelevant features which otherwise lower the classification accuracy.

Tables 7.4 and 7.5 show best results reported in literature in terms of Mean Subset size (denoted as MSS) and classification accuracy (denoted as CA). Further, It is evident from the Tables 1.1, 7.4 and 7.5 that the performance of AGA-MO is better than that of state-of-the-art methods suggested in literature for the dataset *Cleveland*, *Glass* and *Ionosphere*. While using fuzzy rough dependency measure as fitness function, these method provides better accuracy for *Ecoli* and *Wine* dataset at the cost of relatively large reduct size than the existing best method reported in literature. AGA-MO produces the stable result with zero or very low standard deviation.

Table 7.6 shows the results for Friedman test giving the ranking of the performance of AGA-MO, BO and eGA. It is observed that AGA-MO ranks the best as compared to BO and eGA.

Table 7.4: AGA-MO with L-FRFS measure: Comparison of reduct size with statistical t-test, and Wilcoxon test. [Statistical t-test is denoted as 'S'. Wilcoxon test is denoted as 'WT'. Mean Subset Size is denoted as 'MSS'.]

Dataset (Total features)	Feature Selection Method	Feature Subset size				Best MSS reported in literature
		Best	Mean(s.d.)	S	WT	
Cleveland(13)	eGA	6	6(0)	-	-	7.81 [30]
	BO	6	6(0)	-	-	
	AGA-MO	**6**	6(0)			
Ecoli(7)	eGA	5	5(0)	-	-	3 [52]
	BO	5	5(0)	-	-	
	AGA-MO	**5**	5(0)			
Glass(9)	eGA	8	8(0)	-	-	8.44 [30]
	BO	8	8(0)	-	-	
	AGA-MO	**8**	8(0)			
Ionosphere(34)	eGA	6	6.25(0.45)	v	v	7.3 [30]
	BO	6	6.25(0.45)	v	v	
	AGA-MO	**6**	6(0)			
Lung(56)	eGA	4	4.33(0.49)	v	v	NA
	BO	3	4.12(0.16)	v	v	
	AGA-MO	**3**	3(0)			
Soybean small(35)	eGA	2	2(0)	-	-	2 [29]
	BO	2	2(0)	-	-	
	AGA-MO	**2**	2(0)			
Wine(13)	eGA	4	4(0)	-	-	2 [52]
	BO	4	4(0)	-	-	
	AGA-MO	**4**	4(0)			
LSVT(310)	eGA	13	11.41(1.24)	v	v	NA
	BO	7	8.5(1.24)	v	v	
	AGA-MO	**5**	**5.08(0.28)**			

Table 7.5: AGA-MO with L-FRFS measure: Comparison of classification accuracy with statistical t-test, and Wilcoxon test. [Classification accuracy is denoted as 'CA'. Statistical t-test is denoted as 'S'. Wilcoxon test is denoted as 'WT'.]

Dataset (Total features)	Feature Selection Method	Classifier: J48				Classifier: JRip				Classifier: PART				Best CA reported in literature
		Best	Mean	S	WT	Best	Mean	S	WT	Best	Mean	S	WT	
Cleveland(13)	eGA	52.47	52.47(0)	-	-	53.79	53.79(0)	-	-	52.14	52.14(0)	-	-	
	BO	52.47	52.47(0)	-	-	53.79	53.79(0)	-	-	52.14	52.14(0)	-	-	52.6 [30]
	AGA-MO	52.47	52.47(0)			**53.79**	53.79(0)			52.14	52.14(0)			
Ecoli(7)	eGA	82.44	82.44(0)	-	-	81.25	81.25(0)	-	-	80.65	80.65(0)	-	-	
	BO	82.44	82.44(0)	-	-	81.25	81.25(0)	-	-	80.65	80.65(0)	-	-	77.38 [52]
	AGA-MO	**82.44**	82.44(0)			81.25	81.25(0)			80.65	80.65(0)			
Glass(9)	eGA	64.49	64.49(0)	-	-	69.16	69.16(0)	-	-	68.69	68.69(0)	-	-	
	BO	64.49	64.49(0)	-	-	69.16	69.16(0)	-	-	68.69	68.69(0)	-	-	65.14 [30]
	AGA-MO	64.49	64.49(0)			**69.16**	69.16(0)			68.69	68.69(0)			
Ionosphere(34)	eGA	92.6	89.84(1.76)	-	-	92.59	89.48(1.61)	-	v	**93.45**	89.71(1.75)	-	*	
	BO	91.45	88.55(1.12)	-	-	92.31	90.1(1.1)	-	-	91.17	89.8(0.99)	-	-	86.17 [30]
	AGA-MO	91.45	89.58(1.15)			92.31	90.33(1.04)			91.17	89.5(0.81)			
Lung(56)	eGA	87.5	75.51(7.75)	v	v	87.5	74.73(9.55)	v	v	87.5	76.04(7.57)	v	v	
	BO	87.5	83.67(1.61)	-	-	87.5	83.67(1.51)	-	-	87.5	81.85(2.6)	-	-	NA
	AGA-MO	**87.5**	85.67(1.61)			87.5	85.67(1.61)			87.5	83.85(2.61)			
Soybean small(35)	eGA	100	100(0)	-	-	100	100(0)	-	-	100	100(0)	-	-	
	BO	100	100(0)	-	-	100	100(0)	-	-	100	100(0)	-	-	100 [29]
	AGA-MO	**100**	100(0)			100	100(0)			100	100(0)			
Wine(13)	eGA	93.82	92.51(0.67)	v	v	92.13	91.57(0.83)	*	*	93.82	92.08(1.00)	v	v	
	BO	93.82	93.82(0)	-	-	89.89	89.89(0)	-	-	93.82	93.82(0)	-	-	90.44 [52]
	AGA-MO	**93.82**	93.82(0)			89.89	89.89(0)			93.82	93.82(0)			
LSVT(310)	eGA	80.16	73.54(4.05)	v	v	**82.54**	73.27(4.91)	v	v	80.95	72.15(5.09)	v	v	
	BO	80.16	74.28(2.27)	-	-	80.16	75.25(2.13)	-	-	81.74	76.18(3.01)	-	-	NA
	AGA-MO	80.16	76.98(2.77)			80.16	**77.25(2.53)**			81.74	77.18(3.01)			

112

Table 7.6: Friedman ranking with L-FRFS measure

Methods	FR	Rank
AGA-MO	1.7031	1
BO	2.0000	2
eGA	2.2969	3

7.3   Conclusion

A new evolutionary optimization algorithm, AGA-MO, has been proposed in this chapter. The proposed AGA-MO optimization method has been applied for feature selection using RST and L-FRFS measures. AGA-MO utilizes DS-initialization and gives better reducts without compromising with classification accuracy. The Superiority of AGA-MO over BO and eGA has been established in terms of reduct size, classification accuracy, t-test, Wilcoxon test and Friedman test. Results of the parametric test (t-test) and non-parametric test (Wilcoxon test) have been discussed and from these results, it is observed that in general the AGA-MO has acceptable classification accuracy.

Further, AGA-MO has the edge, i.e. its performance is better than BO and eGA, in case of RST as well as L-FRFS measures, as suggested by Friedman ranking test. The AGA-MO has shown its effectiveness on large and practical datasets where feature selection has significance.