Immune Grey Wolf Optimizer for Attribute Reduction: Application for Medical Systems

Badra Khellat Kihel[1,2]*   **Souad Khellat Kihel[1]**   **Samira Chouraqui[1]**

[1]*Department of Informatics, Faculty of Mathematic and Informatics,*
Université des Sciences et de la Technologie d'Oran Mohamed-Boudiaf USTO-MB,
B.P. 1505 EL M'NAOUER 31000 Oran, Algeria
[2]*Department of Economics, Oran2 University Oran, Algeria*
* Corresponding author's Email: badra.khellat@univ-usto.dz

Abstract: The sizes of medical datasets are increased with an exponential rate. Extract useful knowledge from those data, analyzing and searching on this very vast amount of data is become a hard task for the scientists. Several metaheuristic algorithms for solving such problem were proposed for dimensionality reduction. In this work, a new stochastic search strategy inspired by the Grey Wolf optimization theory is proposed for feature subset selection. Grey Wolf Optimization algorithm (GWO) is a new metaheuristic optimization technique. Its principle is to reproduce the behavior of grey wolves in nature to hunt in a cooperative way. In order to improve the disadvantage of Grey Wolf Optimizer when solving the feature selection problem, three immune operators (selection, cloning and mutation) are embedded into the standard GWO and an Hybrid Immune Grey Wolf Algorithm (IGWO) is proposed. To ensure the diversity of the population of wolves and avoid premature convergence, we cloned and mutated Alpha, Beta and Delta wolves before selection of the new first, second, and third best solutions. The experiments were performed using ten medical datasets: Parkinson, Leukemia and heart diseases, Breast, Colon, and Prostate Cancers, and the results show that the proposed approach is able to explore different regions of a search space and avoid local optima reducing the dimensionality of datasets than can significantly improve the performances of the system for diseases recognition and decrease the required classification time.

Keywords: Grey wolf optimizer, Artificial immune system, Clonal selection, Medical system, Feature selection.

1. Introduction

In a computer aided diagnosis system, there is always a large amount of data which must be all saved and used in learning techniques to classify healthy and pathological patients. Achieving a successful diagnostic process must take into account several considerations related to knowledge extraction, compact representation and data analysis.

The problem of dimensionality reduction in medical system has been widely investigating due to its importance. Feature selection (FS) allows the reduction of feature space, which is crucial in reducing the training time and improving the prediction accuracy. This is achieved by removing irrelevant, redundant, and noisy features (i.e. selecting the subset of features that can achieve the best performance in terms of accuracy and computational time) [1, 2].

Many scholars have researched on feature selection methods. Two main methods for feature selection can be distinguished: Wrapper approaches include a learning-classification algorithm in the evaluation procedure, while filter approaches do not include such algorithm. Filter approaches are argued to be computationally less expensive and more general, while wrapper approaches can usually achieve better results [3].

Recently, an optimization technique called grey wolf optimizer algorithm has gained the interest of many researchers. This algorithm is a type of swarm intelligence algorithm based on the position change

of wolves according to the experience of leaders: Alpha, Beta and Delta [4, 5].

In this way, the wolves will gradually become the same as the leaders. To avoid premature convergence, we proposed a new Immune Grey Wolf Optimizer (IGWO), using immunological operators (cloning, mutation and selection) on Alpha, Beta and Delta wolves to update the wolves' position, to select the best feature subset which include a small number of features and achieve a lower classification error rate than using all available features. The selected methods were evaluated using ten well-known medical datasets. The obtained results are promising and confirm the interest of using bio-inspired approaches for data mining in bioinformatics.

The rest of this paper is organized as follows: section 2 reviews existing works. Section 3 introduces bio-inspired algorithms (GWO and AIS). The proposed Immune grey wolf Optimizer Algorithm is clearly described in section 4. Section 5 illustrates and analyzes the experimental results. Finally, section 6 concludes the paper.

2. Related works

Recently, many FS based on new optimizers were proposed. Chen et al. proposed a novel rough set based method for feature selection using fish swarm algorithm [7]. The fish swarm algorithm is used for optimization through imitating the fish behaviours such as swarming, following and moving randomly. Mirjalili proposed a wrapper-feature selection approach that uses the binary dragonfly algorithm as a search strategy and the K-Nearest Neighborhood (KNN) classifier as an evaluator [8]. Another recent optimizer is Ant Lion that mimics the performances of ant lions in hunting prey is used as a wrapper feature selection method in [9, 10]. Firefly algorithm (FFA) was combined with artificial immune system (IFA) to reduce dimensionality in voluminous datasets in [11] to prevent the population from being trapped into local optimum. Grey Wolf Optimizer (GWO) is a popular swarm intelligence technique that has been successfully used for solving feature selection problems [4, 5]. In [4], GWO is used to find the best features in the Coronary artery disease identification dataset then Support Vector Machine classifier (SVM) for the evaluation of the GWO fitness function.

Grey wolf optimizer (GWO) has received widespread attention from scholars. The algorithm was inspired by grey wolves' predation activity in nature. This algorithm has gained much interest in the hybrid metaheuristics field. In [13] authors proposed a multi-objective grey wolf optimizer in order to optimize problems with multiple objectives with a novel leader selection mechanism. An evolutionary GWO algorithm was proposed in [14] to resolve the distributed generation allocation problem using genetic algorithm. In [15], the authors proposed a hybrid GWO and Artificial Bee Colony (ABC) to improve the complex systems performance. Al-Tashi et al. proposed a binary version of hybrid particle swarm optimization PSO-GWO to solve feature selection problem [16]. In another study PSO was combined with GWO to outperform the convergence speed of the system [17]. Kohli et al. introduce the chaos theory into the GWO for the purpose of accelerating the global convergence speed [18]. These studies have revealed that the hybrid algorithms performed much better compared to standard metaheuristic methods.

In this work, we were particularly attracted by the use of bio-inspired methods and there hybridization for dimensionality reduction in medical systems. We have used two algorithms: Grey Wolf Optimizer and Artificial Immune System (AIS) to select the most relevant features in medical datasets, then we produce a new hybrid Immune Grey Wolf Optimizer (IGWO) in which we used GWO to select the three best wolves: Alpha, Beta and Delta then we used AIS in particular cloning, mutation and selection operators for the three best wolves to increase the global search space. The main aim of this hybridization is to study the influence of the initial population quality on the searching progress of GWO in feature selection task.

3. Methods

3.1 Grey Wolf Optimizer (GWO)

Grey Wolf Optimizer is a recent metaheuristic, which offers better performance in feature selection task. However, the new positions of wolves are mostly based on the experience of three solutions or leaders: Alpha, Beta and Delta in the position update [4].

In the feature selection problem, a representation for candidate feature subset must be chosen and encoded. In most studies, wolf is a binary string of length equal to the total number of features so that each bit encodes a single feature. A bit of '1' ('0') implies the corresponding feature is selected (excluded).

For modeling the social hierarchy of wolves, the leaders of the park are considered as the alpha (α). The main responsibility of alpha is making decisions.

Beta (β) known to assist alpha in making decisions and the main responsibility of beta is the feedback suggestions. Delta (δ) performs as scouts, and controls omega (ω) wolves by obeying alpha and beta wolves. The omega wolves must obey every other wolf.

In GWO, α, β, and δ, guides the hunting process and ω wolves follows them. The encircling behavior for the pack to hunt a prey can be expressed as

$$X(t + 1) = X_p(t) - A.D \tag{1}$$

Where X_p is the position of prey, D is defined as:

$$D = |C.X_p(t) - X(t)| \tag{2}$$

Where X is the position of grey wolf, and t is the number of iterations. A and C are coefficient vectors calculated as follows:

$$A = 2a.r_1 - a \tag{3}$$

$$C = 2.r_2 \tag{4}$$

Where $r1$ and $r2$ are independent random numbers between [0,1], and a is the encircling coefficient that is linearly decreased, from 2 to 0.

The leaders are guiding the omega wolves to move toward the optimal position. Therefore, we save the first three best solutions obtained and force the other search agents to update their positions according to the position of the best search agents.

$$D_\alpha = |C_1.X_\alpha - X| \tag{5}$$
$$D_\beta = |C_2.X_\beta - X| \tag{6}$$
$$D_\delta = |C_3.X_\delta - X| \tag{7}$$

Mathematically, the new position of wolf is:

$$X(t + 1) = \frac{X_1 + X_2 + X_3}{3} \tag{8}$$

Where X_1, X_2 and X_3 are the three best wolves in the swarm at a given iteration t calculated as follows:

$$X_1 = |X_\alpha - A_1.D_\alpha| \tag{9}$$

$$X_2 = |X_\beta - A_2.D_\beta| \tag{10}$$

$$X_3 = |X_\delta - A_3.D_\delta| \tag{11}$$

Fig. 2 shows how a search agent updates its position according to alpha, beta, and delta in a search space.
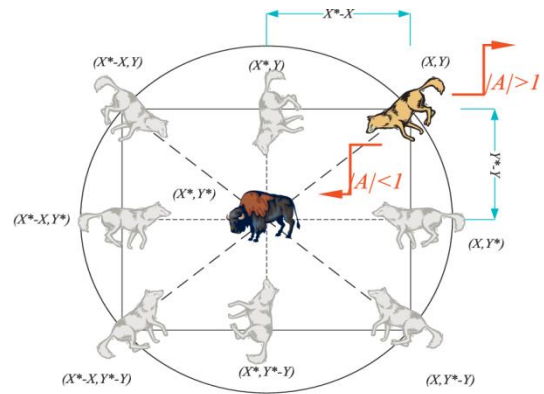

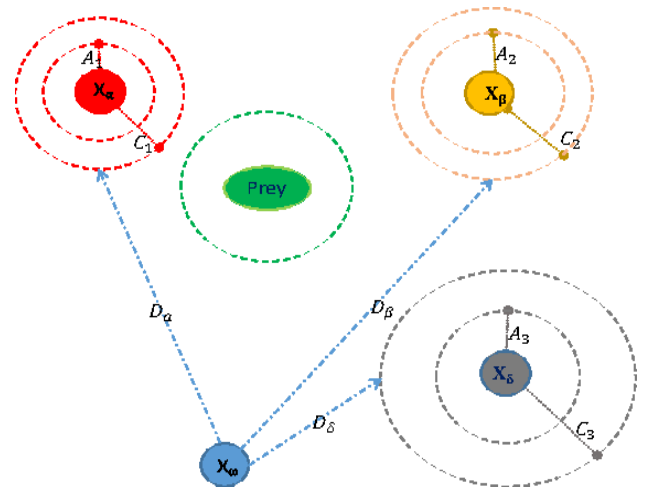Figure. 1 Position updating mechanism of search agents and effects of A on it [19]

Fig. 2 Position updating in GWO

3.2 Artificial Immune System (AIS)

The artificial immune systems work on three basic immunological principles which include clonal selection theory, negative selection principles and the mechanism of immune network [11, 20]. The main idea of the Clonal selection method is to multiply only the cells whose antibodies are able to recognize the antigens [21].

In order to clarify how an immune response is mounted when a nonself antigenic pattern is recognized by a B cell, the clonal selection theory has been developed. The clonal selection algorithm can be described as follows:

Step 1: Generate a set P of candidate solutions, composed of the subset of memory cells M added to the remaining population Pr.

Step 2: Select the n best individuals of the population, based on an affinity measure.

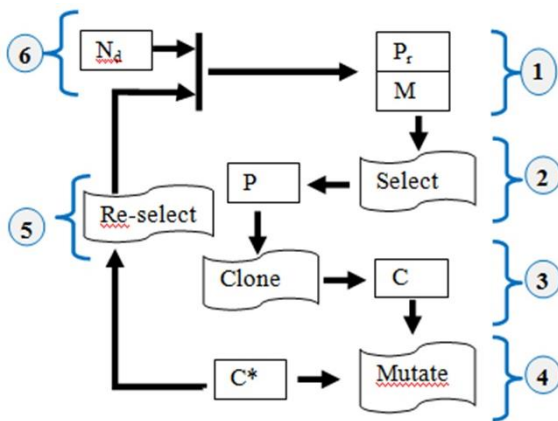$$Affinity = max\{accuracy\ (sub)\} \tag{12}$$

Figure. 3 Clonal Selection Algorithm

Where *accuracy (sub)* evaluates the classification accuracy of the subset.

Step 3: Clone the *n* best individuals of the population to a temporary population of clones *C* according to their affinity. The number of clones for each antibody *i* is computed by

$$NC(i) = round\left(\frac{B.affinity_i^2}{\sum_j affinity_j^2}\right) \qquad (13)$$

Where *B* is the cloning parameter, affinity is calculated in Eq. (12) and *round ( )* is the operator that rounds its argument to closest integer.

Step 4: A maturated antibody population is generated C^*.

The mutation changes according to the representation of the data; in this direction we find various types of mutation in the case of real representation or binary presentation (which is our case).

In the first step we must calculate the number of bits to be inversed *(Nm)* using:

$$Nm = round\left(\left(L - affinityC(i)\right).rand()\right) \qquad (14)$$

Where *L* represents the antibody size and *affinityC* is the vector of affinity calculated after cloning. *Rand ( )* is a mathematical function used to generate a random number that is greater than or equal to 0 and less than 1.

In binary valued individuals mutation is the flipping of variable values (inverse '0' to '1' and vice versa),

Step 5: Re-select the improved individuals from C^* to compose the memory set M. Some members of *Ab* can be replaced by other improved members of C^*.

Step 6: Replace *d* antibodies by new ones (diversity introduction). The lower affinity cells have higher probabilities of being replaced [11, 20, 21].

4. Hybrid system IGWO

Several bio-inspired hybrid methods were proposed to perform FS problems. In this context, after a detailed study of bio-inspired approaches, we propose immunological and swarm intelligence methods for feature selection. Our study of immunological approaches oriented us to propose the Artificial Immune System (AIS) and our study of swarm intelligence methods guided us to opt Grey Wolf Optimizer (GWO) then their hybridization (IGWO).

In this work, we have used Grey Wolf Optimizer and Artificial Immune System to select the most relevant features with the use of k-nearest neighbors in each iteration to evaluate the error rate of the selected features.

The IGWO's basic idea is to increase the algorithm's capability to exploit AIS with the ability to explore GWO to achieve good results. Initially, the population of grey wolves is randomly initialized (either bit 1 or 0). Afterward, the fitness of each wolf is evaluated using Knn. The best, second, and third best solutions defined as alpha, beta, and delta will be cloned then mutate using Eq.13 and Eq.14. Then, the position of wolf is updated by using Eq.8 on the selected three best wolves X_α, X_β and X_δ from the mutated population.

Next, the fitness of each wolf is evaluated and the positions of alpha, beta, and delta are updated. The algorithm is repeated until the stopping criterion is satisfied (number of iterations). Finally, the alpha solution is selected as the optimal feature subset.

5. Experimental results

5.1 Medical datasets used

The proposed system was evaluated on ten medical datasets: Parkinson, Leukemia and heart diseases, Breast, Colon, and Prostate Cancers.

5.2 Parameters tuning

In order to approve our system, few parameters tuning has to be made. Several parameters for GWO and AIS are required to be tuned to find the best value that can produce the optimum or near optimum.
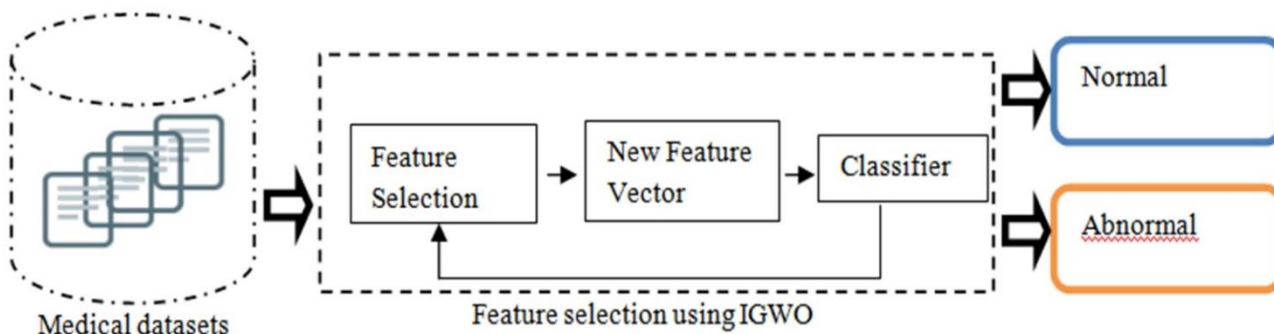
Figure. 4 Training phase of proposed hybrid algorithm IGWO

Table 1. Datasets used

Data set	Number of Features	Number of Classes	Number of Instances
Spect Heart (Binary)	23	02	187
Spect Heart	44	02	187
Parkinson'Diseace	22	02	96
Parkinson 2	26	02	1040
Breast tissue	09	06	106
Wisconsin Breast Cancer	09	02	699
Wisconsin Breast Cancer Diagnostic	31	02	569
Colon	2000	02	62
Leukemia	7070	02	72
Prostate	5966	02	102

The population size of grey wolves is taken 10 and the maximum number of iterations is fixed at 100 iterations. The values of two vectors r1 and r2 are taken randomly in the range of (0, 1) and the controlling parameter a has linearly decreasing values from 2 to 0 over the course of iterations.

For AIS, the cloning parameter, b, is set at 0.1, the population size, N, and maximum number of iterations, T, are fixed at 10 and 100, respectively.

5.3 Results and discussion

In this section we present the experimental results obtained for the different approaches on datasets. Table 2 presents the results of the IGWO algorithm for feature selection in terms of number of selected features, classification error and CPU time. The IGWO is run 10 times on the ten medical databases and the statistics results of these 10 runs are provided Table 2.

In addition, the IGWO algorithm is compared with AIS and GWO algorithms in Table 3

The results show that the proposed method is able to produce good performance on reducing the effects of the outliers and the noises and improve classification accuracy for disease recognition. The features selected can accomplish the goal of achieving higher accuracy with smaller size of features.

The results of the proposed algorithm show that the use of immunological operators for Alpha, Beta and Delta wolves influence the robustness and the algorithm convergence, by using the cloning, mutation and selection operators to calculate position that ensures diversity of the population

Fig.6 shows the Box-plot of error rates and number of selected features of the used feature selection methods for different datasets.

Metaheuristic methods have the limitation that, they often need tuning on the part of the user, to work correctly. Also, being random in nature, the solution produced is often nearly the global optimum. IGWO performs so well that, the results produced by it match the optimal by exhaustive search.

5.4 Comparative study

In this section, the results of the conducted experiments on the proposed systems are compared to some existed works in the literature. Table 4 depicts the error classification rate (Err) and the number of selected features (NF) from the proposed system as well as the existing swarm. The results showed a good performance of the proposed approach comparing to the existing algorithms. This is mainly due to the high exploration of the IGWO, because of the position for updating equations of GWO when combining the three best solutions Alpha, Beta and Delta.

6. Conclusion

In this work, a dimensionality reduction system has been evaluated on medical systems for the aim of disease recognition approving. Two metaheuristic algorithms have been studied such as Artificial
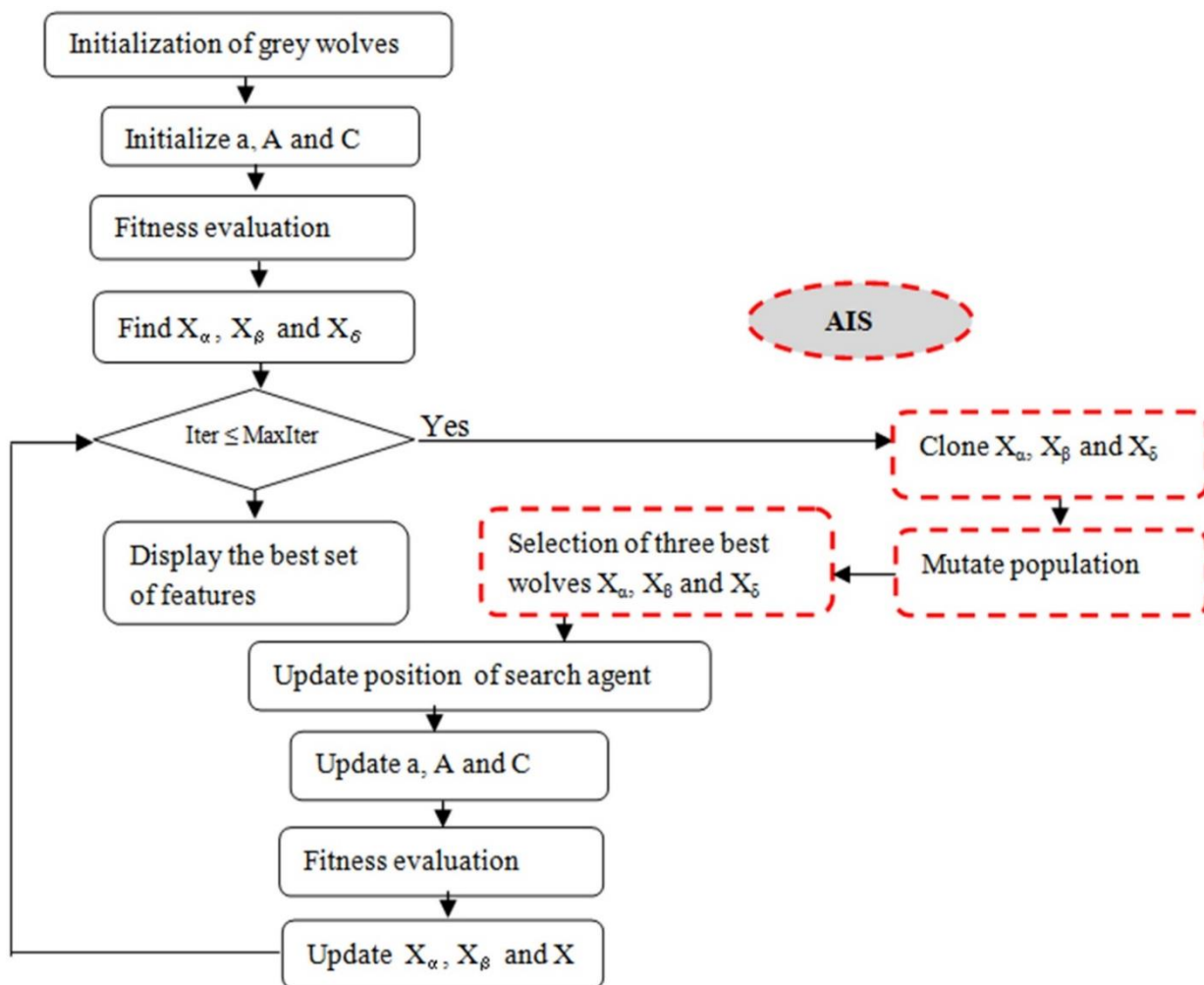
Figure. 5 Immune Grey Wolf Optimizer Algorithm

Table 2. Results for hybrid IGWO

Data set	Number of Features (NF)				Error rate (Err)				CPU time
	Best	Worst	Average	Std	Best	Worst	Average	Std	
Dataset1	11	14	12.2	1.303	0.15	0.2	0.18	0.022	1023.27
Dataset2	12	20	15.6	4.037	0.1125	0.2	0.152	0.035	1035.16
Dataset3	8	10	9	1	0.0416	0.062	0.047	0.009	935.69
Dataset4	7	13	10.4	2.190	0.3192	0.335	0.326	0.006	1168.,79
Dataset5	4	6	4.8	0.836	0.3207	0.349	0.335	0.010	1064.01
Dataset6	6	7	5.8	0.836	0.0200	0.025	0.022	0.002	1151.34
Dataset7	9	13	10.4	1.949	0.0421	0.045	0.043	0.001	1099.49
Dataset8	823	1011	880.6	77.125	0.0967	0.145	0.125	0.021	1107.36
Dataset9	2520	3132	2762.2	257.740	0.0555	0.069	0.058	0.006	1177.38
Dataset10	2124	2530	2359	172.594	0.0882	0.107	0.103	0.008	1284.42

Immune System (AIS) and Grey Wolf Optimizer (GWO). Then, a new hybrid variant is proposed by using strengths of GWO and AIS.

The proposed algorithm was tested over ten different medical datasets: Parkinson, Leukemia and heart diseases, Breast, Colon, and Prostate Cancers. The obtained results after reduction are very encouraging and prove that hybrid approach is more reliable in giving best results as compared to some bio-inspired algorithms: AIS, GWO, ABC, ACO, FFA and IFA.

As future work, the proposed algorithm can be experimented with other classifiers such as Support Vector Machine (SVM) or Artificial Neural Network (ANN).

Table 3. Results for algorithms used (GWO, AIS, IGWO)

| Data set	GWO		AIS		IGWO	
	NF	Err	NF	Err	NF	Error
Dataset1	10	0.325	9	0.212	11	0.15
Dataset2	16	0.150	21	0.187	12	0.112
Dataset3	11	0.020	12	0.083	8	0.041
Dataset4	14	0.341	13	0.337	7	0.319
Dataset5	04	0.264	05	0.311	4	0.320
Dataset6	06	0.025	07	0.027	6	0.020
Dataset7	07	0.042	16	0.049	9	0.042
Dataset8	840	0.145	1022	0.096	823	0.096
Dataset9	3150	0.069	3547	0.055	2520	0.055
Dataset10	2552	0.078	2951	0.107	2124	0.088
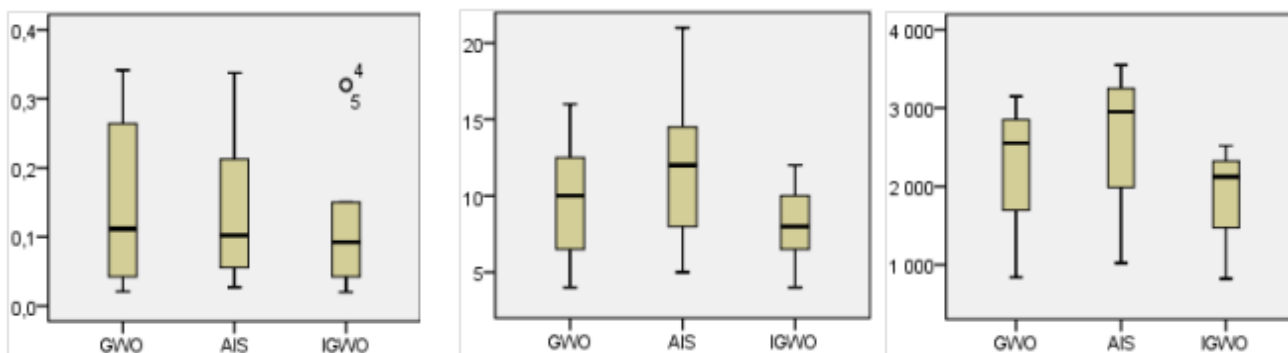
Figure. 6 Box-plot for feature selection methods (a) Box-plot of error rate (b) Box-plot of number of selected features for small datasets and (c) for big datasets.
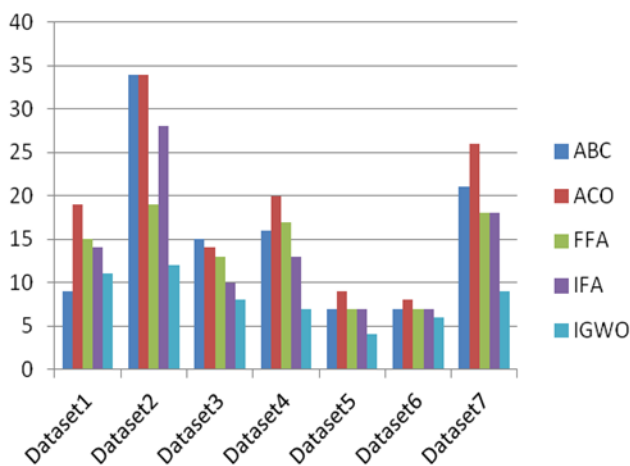

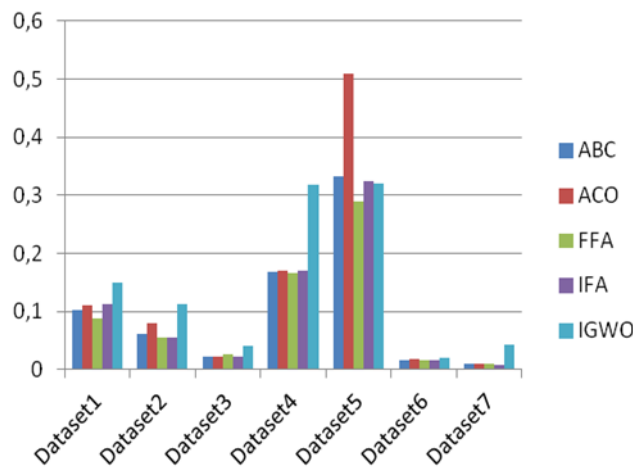Figure.7 Number of selected features for comparative study

Figure.8 Error rates for comparative study

Table 4. Comparative study

Data set	ABC [22]		ACO [22]		FFA [11]		IFA [11]		IGWO	
	NF	Err	NF	Err	NF	Err	NF	Err	NF	Err
Dataset1	9	0.102	19	0.111	15	0.088	14	0.112	11	0.15
Dataset2	34	0.062	34	0.080	19	0.056	28	0.056	12	0.112
Dataset3	15	0.022	14	0.023	13	0.027	10	0.023	8	0.041
Dataset4	16	0.169	20	0.170	17	0.166	13	0.171	7	0.319
Dataset5	7	0.332	9	0.510	7	0.290	7	0.325	4	0.320
Dataset6	7	0.017	8	0.019	7	0.016	7	0.017	6	0.020
Dataset7	21	0.010	26	0.010	18	0.009	18	0.008	9	0.042

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

The paper conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing-original draft preparation, writing-review and editing, visualization, have been done by 1st and 2nd authors.

The supervision and project administration have been done by 3rd author.

References

[1] B. Xue, "Particle Swarm Optimization for Feature Selection in Classification: A Multi-Objective Approach", *IEEE Transactions on Cybernetics*, Vol. 43, No. 6, pp. 1656–1671, 2013.

[2] H. Liu and L. Yu, "Toward Integrating Feature Selection Algorithms for Classification and Clustering", *Knowledge and Data Engineering*, Vol. 17, No. 4, pp. 491-502,2005.

[3] T. Mohammed , S.Alhayali , O.Bayat and O. Uçan, "Feature Reduction Based on Hybrid Efficient Weighted Gene Genetic Algorithms with Artificial Neural Network for Machine Learning Problems in the Big Data", *Scientific Programming*, Vol. 2018, 2018.

[4] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer", *Advances in Engineering Software*, Vol. 69, pp. 46-61, 2014.

[5] Q. Al-Tashi, H. Rais, and S. Jadid, "Feature Selection Method Based on Grey Wolf Optimization for Coronary Artery Disease Classification", In: *Proc. of International Conference of Reliable Information and Communication Technology*, Kuala Lumpur, Malaysia, pp. 257–266, 2018.

[6] M. Esseghir, "Metaheuristics for The Feature Selection Problem : Memetic, Adaptive and Swarm Approaches", *Thèse de Doctorat Université d'Artois*, 2011.

[7] Y. Chen, Q. Zhu, H. Xu; "Finding rough set reducts with fish swarm algorithm", *Knowl-Based Syst.*, Vol. 81, 22–29, 2015.

[8] S. Mirjalili, "Dragonfly algorithm: a new metaheuristic optimization technique for solving single objective, discrete, and multi-objective problems", *Neural Computing and Applications*, Vol. 27, pp. 1053- 1073, 2016.

[9] S. Mirjalili, "Hybrid Binary Ant Lion Optimizer with Rough Set and Approximate Entropy Reducts for Feature Selection", *Soft Computing*, pp. 1-17, 2018.

[10] E. Emary and H. M. Zawbaa, "Feature selection via Levy Antlion optimization", *Pattern Analysis and Applications*, Vol. 21, pp. 1–20, 2018.

[11] B. K. Kihel and S. Chouraqui, "Firefly Optimization Using Artificial Immune System for Feature Subset Selection", *International Journal of Intelligent Engineering and Systems*, Vol. 12, No. 4, pp. 377-347, 2019.

[12] P. Anand and S. Arora, "A novel chaotic selfish herd optimizer for global optimization and feature selection", *Artificial Intelligence Review*, Vol. 51, pp. 1-46, 2019

[13] S. Mirjalili, S. Saremi, S.M. Mirjalili,L. S.Coelho, "Multi-objective grey wolf optimizer : A novel algorithm for multi-criterion optimization", *Expert Systems with Applications*, Vol. 47, pp. 106-119, 2016.

[14] R. Sanjay, T. Jayabarathi,T. Raghunathan, V. Ramesh, and N. Mithulananthan, "Optimal Allocation of Distributed GenerationUsing Hybrid Grey Wolf Optimizer", *Access IEEE*, Vol. 5, pp. 14807-14818, 2017.

[15] P. J. Gaidhane and M. J. Nigam, "A Hybrid Grey Wolf Optimizer and Artificial Bee Colony Algorithm for Enhancing the Performance of Complex Systems", *J. Comput. Science*, 2018.

[16] Q. Al-Tashi, S. Abdulkadir, H. Rais, S. Mirjalili, and H. Alhussian," Binary Optimization Using Hybrid Grey Wolf Optimization for Feature Selection", *Access IEEE*, Vol. 7, pp. 39496-39508, 2019

[17] N. Singh, S.B. Singh," Hybrid Algorithm of Particle Swarm Optimization and Grey Wolf Optimizer for Improving Convergence Performance", *Journal of Applied Mathematics*, Vol. 2017, 2017.

[18] M. Kohli and S.Arora, "Chaotic grey wolf optimization algorithm for constrained optimization problems", *Journal of Computational Design and Engineering*, Vol. 5, pp. 458-472, 2018

[19] M. Majid, G. Milad, S. Hossein, M. Jalal and G. Gevork, "Application of multi-objective grey wolf algorithm on energy management of microgrids with techno-economic and environmental considerations" , In: *Proc. of 3rd Conference on Swarm Intelligence and Evolutionary Computation,* Higher Education Complex of Bam, Iran, 2018.

[20] L. N. De Castro and F. J. Von Zuben, "Learning and optimization using the clonal selection

principle", *IEEE Transactions on Evolutionary Computation*, Vol. 6, pp. 251-239, .2002.

[21] L. N. De Castro and F. J. Von Zuben, "The clonal selection algorithm with engineering applications", In: *Proc. of GECCO*, 2000.

[22] B. Khellat and M. Benyettou, "Hybrid Bio-inspired Approach for Feature Subset Selection", *International Journal of Applied Engineering Research*, Vol. 13, No.10, pp. 7895-7902, 2018.