Received: April 21, 2019

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Firefly Optimization Using Artificial Immune System for Feature Subset Selection

Badra 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,
B.P. 1505 EL M'NAOUER 31000 Oran, Algeria
* Corresponding author's Email: badra.khellat@univ-usto.dz; khellat_badra@yahoo.fr

Abstract: At first glance, the feature selection is a crucial step in a pattern recognition system. The main objective of this selection is to reduce the features number, by eliminating irrelevant and redundant attributes. In addition, we try to maintain or improve the classifier performance using neural network algorithm. Nevertheless, a new stochastic search strategy inspired by the clonal selection theory in an artificial immune system is proposed for feature subset selection. We have used the firefly and clonal selection algorithms to select the most relevant features in a dataset. In our proposed strategy, feature selection algorithm is formulated as an optimization problem that searches an optimum with less number of features in a feature space and a good accuracy. The goal of our study is to achieve a balance between the classification accuracy and the size of the feature subsets selected using two new hybrid algorithms based on Immune Firefly Algorithm (IFA). Our proposed approach has been evaluated on 10 standard datasets taken from UCI repository. The experimental outcomes have been compared to some popular feature selection methods. The comparison of results shows that our methods significantly outperform most of the used feature selection algorithms.

Keywords: Artificial immune system, Clonal selection algorithm, Feature selection, Firefly algorithm, Neural network.

1. Introduction

Optimization problem has been a hard task to many researchers in order to find the best local searching method. This problem also leads to a branch of knowledge which is the evolutionary computing. The methods were greatly influenced by nature. Few decades ago, many methods have been developed, for instance, Particle Swarm Optimization (PSO), firefly Algorithm (FFA) or Artificial Immune System (AIS). In this study, the hybrid artificial immune system-firefly algorithm is evaluated in comparison with traditional evolutionary algorithms [1].

Feature selection in a high-dimensional data space can decrease the computational cost and may also improve the accuracy during the classification process. Given a set of measurements, reduction of dimensionality can be achieved in essentially two ways:

- Feature extraction is to find the transformation from a higher to a lower dimensional feature space with most of the desired information content preserved.
- Feature selection is used to identify the variables that do not contribute to the classification process. In a discrimination problem, those variables that do not contribute to class separability will be neglected. Thus, the task of feature selection is to select a subset from a large number of features or variables used in classification while maintaining acceptable classification accuracy [2].

Different methods have been developed and used for feature subset selection using several search strategies and evaluation functions. In [3] a correlation measure is applied to evaluate the goodness of feature subsets based on the hypothesis that a good feature subset is one that contains features highly correlated to the class, yet uncorrelated to each other.

Recently developed nature-inspired optimization algorithms are the best approaches for finding global solutions for combinatorial optimization problems like microarray datasets. Different genetic algorithm (GA) methods have been proposed to tackle the feature selection [5, 6]. A hybrid algorithm was proposed using GA with the artificial neural networks (ANN) where GA was used as a pre-step to reduce the feature size [7]. In [8] Al-ani proposed an Ant Colony Optimization (ACO) approach to solve feature selection (FS) problem. His iterative algorithm starts by the selection of random starting point for each ant (initial feature added to the solution subset), and then used pheromone to guide network exploration to make a final subset of features. Khushaba has proposed a hybrid system based on ACO and Differential Evolution (DE). The DE crossover and mutation were applied at the end of each iteration. The resulting subsets are then used to update pheromone trails and the process restarts [9]. Based on Particle Swarm Optimization PSO, Unler proposed a feature selection algorithm with an adaptive selection strategy, where he used the features already selected to select a new feature so a feature is chosen not only according to the likelihood calculated by PSO but also to its contribution to the features already selected [10]. In [4] Liu proposed a Fast Wrapper Feature Subset Selection method based on Binary Particle Swarm Optimization (BPSOWFSS), in which the search process of PSO is optimized by using domain knowledge of feature subset selection problems. Artificial Immune System (AIS) was proposed in [11-13] to select relevant features in different domain. In [11], Long-Short-Term Memory (LSTM) recurrent neural networks were trained with the Artificial Immune Recognition System (AIRS) in order to obtain a long-lived unit for selecting features. In [12] AIS was used for short-term electrical load forecasting. In [13] Clonal Selection Algorithm (CSA) was used to select the most excellent weights for every feature in web pages.

Recently, many FS based on new optimizers were proposed in the literature including Grey Wolf Optimizer (GWO) [14, 15], Whale Optimization Algorithm (WOA) [16], Butterfly Optimization Algorithm (BOA) [17], Dragonfly algorithm (DA) [18], Firefly Algorithm (FFA) [19] and Selfish Herd Optimizer (SHO) [20] that have been successfully employed for solving FS problems.

Another recent algorithm is the Ant Lion Optimizer (ALO). ALO algorithm has been proposed for feature selection in [21, 22]. In [21], two incremental hill-climbing techniques (QuickReduct and CEBARKCC) are hybridized with the Binary Ant Lion Optimizer in a model called HBALO. In another work, ALO was combined with GWO for solving feature selection problem [23].

In this work, we were particularly attracted by the hybridization of bio-inspired methods for feature selection. We produce two Immune Firefly Algorithms IFA1 and IFA2.

In IFA1, we have used the FFA and AIS simultaneously. Firefly Algorithm (FFA) is used to increase the global search mobility of fireflies and Clonal Selection Algorithm can be applied 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 main aim of IFA2 is to study the influence of the quality of the initial population on the searching progress of the AIS algorithm. This hybrid algorithm can use a mixture of the principals of the two evolutionary algorithms to improve the solution quality in solving our feature selection problem. The proposed algorithms are tested on 10 well-known datasets and show a very good performance when comparing to other algorithms in the literature.

The rest of this paper is organized as follows: section 2 reviews feature selection algorithm. Section 3 introduces bio-inspired algorithms: artificial immune system and firefly algorithm. Section 4 presents bio-inspired algorithms for feature selection. In section 5, we present the hybrid proposed approaches. The experimental results obtained are presented and discussed in section 6. In section 7 we compare algorithms proposed with other algorithms. Finally, section 8 concludes the paper.

2. Feature subset selection

The identification of useful and informative attributes for a given dataset, broadly referred to as Feature Selection (FS), is an attractive and challenging research topic for several domains including predictive data mining, pattern recognition, machine learning and information retrieval. One of the fundamental motivations for feature selection is to reduce the dimensionality. In fact, the presence of useless features may not only deteriorate the performance of learning algorithms but also obscure

Received: April 21, 2019

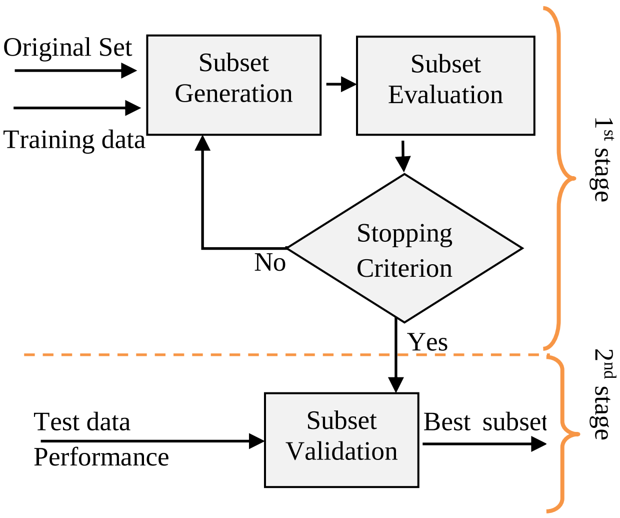

Figure.1 Feature selection process

information behind data. Considered as a fundamental problem in machine learning, the role of FS is critical, especially in a context deemed with irrelevant features [8].

In literature, the authors state a list of three objectives of using feature selection for classification that are:

- Feature extraction reduction task;
- Classification precision improvement;
- Performance estimation reliability improvement [24].

Figure 1 depicts the generic process that could summarize the different steps by any feature selection approach.

As input the feature selection process requires the dataset for which the relevant features will be identified. The outcome should include the retained features.

Generally, such process consists of two stages: search and validation. Each candidate subset is evaluated according to certain criterion and compared to the best solution found. Generation and evaluation are repeated until a given stopping criterion is satisfied. The best subset resulting from the first stage is provided as input for the second stage where it is usually, validated on a different data set [2].

3. Background

3.1 Firefly algorithm optimization

The Optimization problem is one of the most difficult and challenging problems that has received considerable attention over the last decade. Researchers have been constantly investigating better ways to solve it. Recently, one optimization technique called Firefly Algorithm (FFA has gained the interest of many researchers. This algorithm is a type of swarm intelligence algorithm based on the reaction of a firefly to the light of other fireflies [25, 26].

Firefly Algorithm was first developed by Xin-She Yang in late 2007 and 2008 at Cambridge University, which was based on the flashing and behavior of fireflies. In essence, FFA uses the following three idealized rules:

- Fireflies are unisex so that one firefly will be attracted to other fireflies regardless of their sex.
- The attractiveness is proportional to the brightness, and they both decrease as their distance increases. Thus for any two flashing fireflies, the less bright one will move towards the brighter one. If there is no brighter one than a particular firefly, it will move randomly.
- The brightness of a firefly is determined by the landscape of the objective function [19, 27].

According to the above three rules, the degree of attractiveness of a firefly is calculated by the following equation:

$$\beta = \beta_0 \, e^{-\gamma r^2} \tag{1}$$

Where β is the degree of attractiveness of a firefly at a distance r, β_0 is the degree of attractiveness of the firefly at $r = 0$, r is the distance between any two fireflies, and γ is a light absorption coefficient.

The movement of a firefly i is attracted to another more attractive (brighter) firefly j is determined by:

$$x_i^{t+1} = x_i^t + \beta_0 \, e_{ij}^{-\gamma r^2} \left(x_j^t - x_i^t \right) + \alpha_t e_i^t \tag{2}$$

Where the second term is due to the attraction, the third term is randomization with α_t being the randomization parameter. If $\beta_0 = 0$, it becomes a simple random walk [28, 29].

3.2 Artificial immune system optimization

The main goal of the immune system is to protect the human body from the attack of foreign organisms. The immune system is capable of distinguishing between the normal components of our organism and the foreign material that can cause us harm. These foreign organisms are called antigens. The molecules called antibodies play the main role on the immune system response. When an antigen is detected, those antibodies that best recognize an antigen will proliferate by cloning. This process is called clonal selection theory [30].

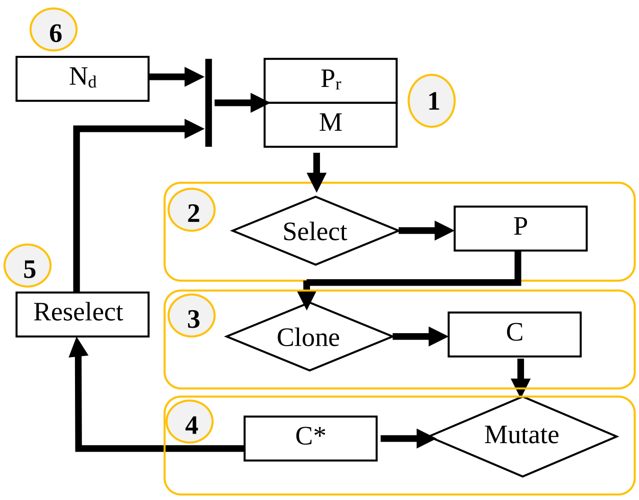

Figure.2 Clonal selection algorithm

3.3 Clonal selection algorithm

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:

(1) Generate a set *P* of candidate solutions, composed of the subset of memory cells *M* added to the remaining population *Pr*.
(2) Select the *n* best individuals of the population, based on an affinity measure.
(3) Clone the *n* best individuals of the population to a temporary population of clones *C*.
(4) Submit the population of clones to a mutation scheme. A maturated antibody population is generated C^*, where the mutation is proportional to the affinity of the antibody with the antigen.
(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^*.
(6) Replace *d* antibodies by novel ones (diversity introduction). The lower affinity cells have higher probabilities of being replaced [30].

4. Feature selection based AIS and FFA

Many methods have been implemented for feature selection and mostly involved statistical approaches. However with the advancement of knowledge and technology, many successful bio-inspired based feature selection algorithms have been proposed.

Banati and Bajaj proposed an algorithm that combines FFA with Rough Set Theory (RST) to ensure the success in less time without compromising the degree of optimality in terms of size of subset and corresponding dependency degree. The algorithm proposed FA_RSAR was evaluated using medical datasets [27]. Another work using FFA for FS was proposed where a simultaneous clustering a feature selection algorithm based on iterative firefly k-means (FKM) algorithm is used. The objective was minimizing the inter-cluster distance as well as maximizing the intra-cluster distance and maximizing the average relevance of the particular feature to the clustering [19]. Sahir and Diri used Long-Short-Term Memory (LSTM) recurrent neural networks which are trained with AIRS in order to obtain the long-lived unit cells for feature selection process [11].

4.1 Feature selection based clonal selection algorithm

The immune system algorithm proposed for feature selection works as follow:

4.1.1. Encoding representation

In the feature selection problem, a representation for candidate feature subset must be chosen and encoded as an antigen. In most studies, an antigen 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).

4.1.2. Affinity calculation

It must find a selective function allowing a good discrimination between the antigens (affinity) to select the best ones then to clone this population and mutate their elements and re-select the best ones for the replacement in the first population if they are better [24]. The objective function is defined by an optimization problem in Eq. (3).

$$Affinity = max\{accuracy\ (sub)\} \tag{3}$$

Where *sub* denotes the corresponding selected feature subset, *accuracy (sub)* evaluates the classification accuracy of the subset.

4.1.3. The cloning

The cloning is the duplication of the data in several specimens; this operation makes possible to keep information long in the workspace. A cloning is proportional to affinity because an antibody approaching more to the antigen is interesting to keep information about it which carry it for a long time and that by duplicating it in several identical specimens, and the mutation will play the role to widen the workspace [31]. We will calculate the number of clones (NC) in our algorithm using:

$$NC(i) = round\left(\frac{B.affinity_i^2}{\sum_j affinity_j^2}\right) \quad (4)$$

Where B is the cloning parameter, affinity is calculated in Eq. (3) and *round ( )* is the operator that rounds its argument to closest integer.

This step raises the average affinity value and gives the following steps a good chance to further move toward the solution. Then, each clone k will be duplicated to $NC(k)$ copies according to Eq. (4).

4.1.4. The mutation

The mutation is defined as an application from Ω to Ω, which associates to each individual X_t a new individual X_{t+1}close to X_t.

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

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

$$Nm = round\ \left(\left(L - affinityC(i)\right).rand()\right) \quad (5)$$

Where L represents the antibody size (the number of features) and *affinityC* in 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), because every variable has only two states (selected (1) or excluded (0)).

4.2 Feature selection based firefly algorithm

Similar to artificial immune system and other evolutionary algorithms, a firefly optimization algorithm is population based approach. A heuristic search using FFA algorithm randomly and uniformly initializes candidate solutions in the parameter space of the problem being solved. During the search process FFA candidates' solution move and collaborate to find the optimal solution to the parameter selection task [19, 28].

4.2.1. Encoding

Consider a feature subset selection problem with n features. Then a feature subset can be represented by a n bits binary string vector $X_i = (x_1, x_2, \ldots, x_n)$ consisted of '0' and '1' If x_k is '0', the k^{th} feature is not selected in this subset. If x_k is '1', the k^{th} feature is selected in this subset. Each binary string vector X_i represents the position of a firefly in FFA.

4.2.2. Representation and updating of attraction and positions

For feature selection, FFA is initialized with a population of N fireflies. Each firefly is treated as a point in an S-dimensional space. The k^{th} firefly is represented as $X_k = (x_1, x_2, \ldots, x_k)$. Each firefly i find its distance with other firefly j and determine the attraction of j with the firefly i.

Each firefly i thus move towards its best mating partner j having minimum distance with i and movement results in greatest increase in dependency. If any firefly that is not able to find any best matting partner, it moves randomly. The movement among fireflies thus results in subsets of fireflies with increase in dependency. The algorithm follows the same procedure for new groups of fireflies generated in previous iteration and determines the accuracy of each group of features selected until the stopping criterion is satisfied (number of iteration) [29].

5. Proposed approach

Recently, another kind of hybridization where the combination is not limited to wrappers and filters or the use of local search to enhance exploitation performance but extended to metaheuristics combination, is explored. Several bio-inspired hybrid methods were proposed to tackle FS problems.

A hybrid algorithm based on swarm intelligence was proposed by combining firefly algorithm and a heuristic method (conditional mutual information maximization). The algorithm proposed is computed in an iterative manner to improve the distribution of information between fireflies and the search efficiency [28]. Another Hybrid Multi-Objective Firefly Simulated Annealing (HMOFSA) algorithm is proposed for Online Feature Selection. The algorithm used the MapReduce paradigm to decompose the original dataset into blocks of examples. Then, HMOFSA algorithm is used to choose the selected features from examples. The attained partial outcomes will be combined into a final vector [29]. Based on swarm intelligence methods: PSO and ACO, Manghour proposed three hybrid bio-inspired approaches for FS task (ACO-PSO1, ACO-PSO2 and ACO-PSO3) [31]. In [32], the Ant Colony, Artificial Bee and Firefly Algorithms were used to select the most relevant features in a dataset then a Genetic Algorithm can create a new population of chromosomes using as initial population the populations generated by the three algorithms used (ACO, ABC and FA) instead of a random one.

In this paper, a feature selection approach based on an Artificial Immune System (AIS) combined with Firefly Algorithm (FFA) is proposed. AIS have the advantage to prevent the population from being trapped into local optimum. Besides, FFA has the ability to improve itself but tend to converge prematurely. Therefore, the combination between FFA and AIS is expected to improve the global search ability and avoid being trapped in local minima even though the population size is relatively small. The idea of our hybrid approaches is to absorb useful information from different feature selection algorithms to find feature subsets that can have smaller size and/or better classification performance than those individual algorithms.

5.1 Immune firefly algorithm IFA 1

Basically, (as shown in fig. 3) the proposed methodology works as follows:

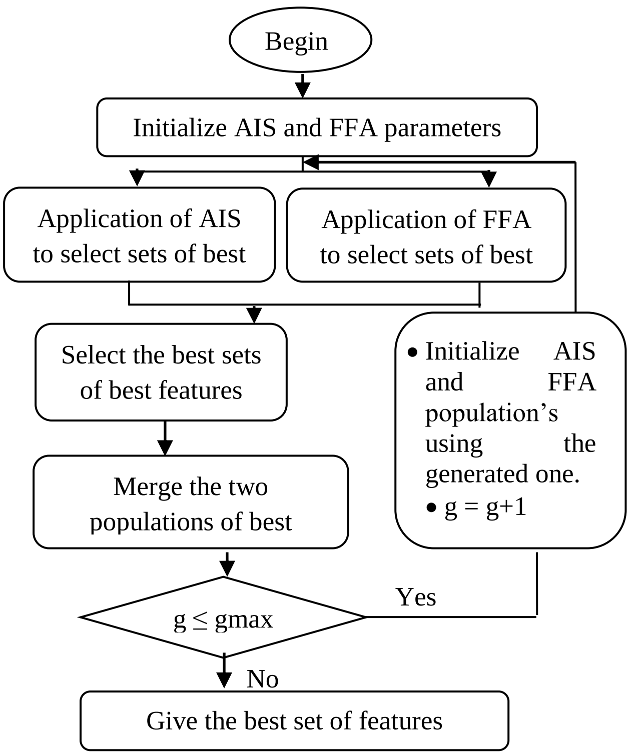

Figure.3 Immune firefly algorithm (IFA1)

Firstly we train each algorithm (AIS and FFA) to produce best sets of features. Then best features of the merged population must be chosen to be used in the system. In the next iteration we will used the chosen population of features for both algorithms (AIS and FFA). The algorithm follows the same procedure for new groups generated in previous iteration and determines the accuracy of each group of features selected until the stopping criterion is satisfied (number of iterations).

The main steps of our algorithm are as follows:

Step 1: Initialization.

- Initialize the FFA parameters and determine the population of fireflies
- Initialize the AIS parameters.
- Determine the maximum of generations (gmax).

Step 2: Run AIS and FFA in the same time

- **FFA**: Generation of fireflies and evaluation of each one.
  - Each firefly *(Fi, i = 1, 2, ..., p)* is randomly assigned to one feature and it should visit all features and build solutions.
  - In this step, the classifier error is used as an evaluation measure.
  - At each iteration, the fireflies build subset using FFA equations Eq. (1), Eq. (2).
- **AIS:** Generation of antibodies and evaluation of each one.
  - Calculate the affinity value for all the population to select best individuals of the population Eq. (3).
  - Operate cloning and mutation to the population, and generate the next generation Eq. (4), Eq. (5).
  - Repeat for some generations.

Step 3:

- Merge the two populations of best features.
- Select best individuals of the population.

- Repeat Steps 2 to 3 until number of generations g=gmax
- Evaluation of the selected subset of features.

5.2 Immune firefly algorithm IFA 2

As shown in fig. 4, we use firefly algorithm (FFA) to fuse multiple feature selection criteria to find the optimal or near optimal subset of informative features and select the most relevant subset of features then we apply the artificial immune system (AIS). With clonal selection algorithm we can create a new population of antibodies using as initial population the population

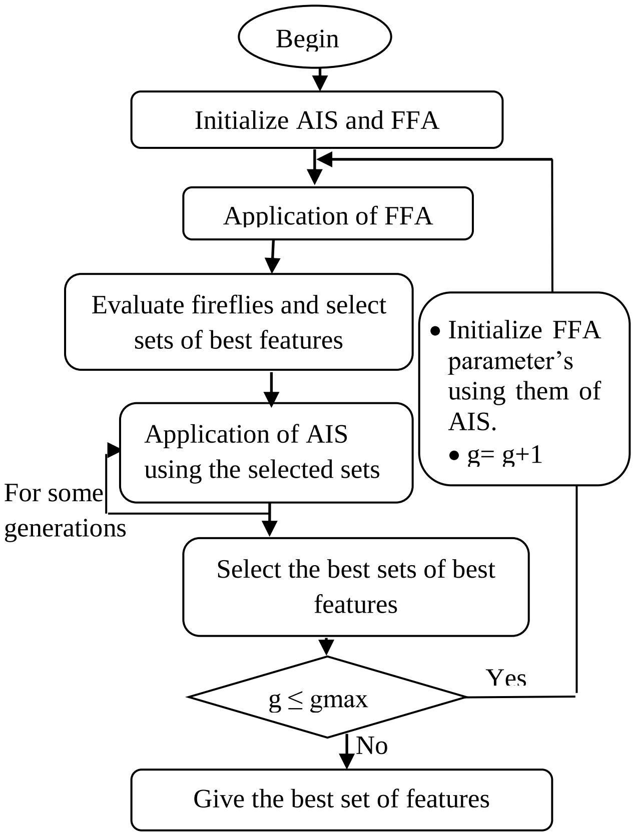

Figure.4 Immune firefly algorithm (IFA2)

generated by FFA instead a random one. A new population is generated by applying immunizing system operators (selection, cloning, and mutation). Our immune system algorithm is designed to maximize classification accuracy and minimize the size of feature subsets. The algorithm follows the same procedure for new groups generated in previous iteration and determines the accuracy of each group of features selected until the stopping criterion is satisfied (number of iterations).

The main steps of our algorithm are as follows:

Step 1: Initialization.

- Initialize the FFA parameters and determine the population of fireflies
- Initialize the AIS parameters.
- Determine the maximum of generations (gmax).

Step 2: Application of FFA

- Generation of fireflies and evaluation of each one.
- Each firefly *(Fi, i = 1, 2, ..., p)* is randomly assigned to one feature and it should visit all features and build solutions.
- In this step, the classifier error is used as an evaluation measure.
- At each iteration, the fireflies build subset using FFA equations Eq. (1), Eq. (2).

Step 3: Application of AIS

- Use the population generated by FFA as initial population of antibodies.
- Calculate the affinity value for all the population to select best individuals of the population using affinity value Eq. (3).
- Operate cloning and mutation to the population, and generate the next generation Eq. (4), Eq. (5).
- Repeat for some generations.
- Select the best sets of best features.

➢ Repeat Steps 2 to 3 until number of generations g=gmax.

➢ Evaluation of the selected subset of features.

6. Experimental results

Initially we describe our dataset as well as the number of features. Then we give the results of the evaluation which relates the number of selected features, the error rates and CPU time.

6.1 Datasets used

In order to make our evaluation results comparable to the most of the published results in feature selection evaluations, we have chosen datasets from the UCI machine learning repository shown in Table 1.

Table 1. Selected datasets (number of features, classes and instances)

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
Glass dataset	10	06	214
Breast tissue	09	06	106
Ionosphere	34	02	351
Musk1	166	02	476
Word Breast Cancer	09	02	699
Word Breast Cancer Diagnostic	31	02	569

Table 2. Results of AIS and FFA feature selection algorithms

Data set	Before selection		AIS			FFA		
	Number of features	Error rate	Number of features	Error rate	CPU time	Number of features	Error rate	CPU time
Spect Heart (Binary)	23	0.1537	11	0.1192	376.70	15	0.0889	382.85
Spect Heart	44	0.1491	18	0.0516	507.82	19	0.0568	458.59
Parkinson's Disease	22	0.0804	13	0.0327	538.71	13	0.0279	495.09
Parkinson 2	26	0.1795	11	0.1778	733.23	17	0.1666	904.17
Glass dataset	10	0.6949	6	0.4963	523.83	7	0.4263	493.17
Breast tissue	9	0.5106	8	0.3215	574.05	7	0.2901	529.70
Ionosphere	34	0.0678	21	0.0340	555.59	19	0.0333	599.52
Musk1	166	0.0516	91	0.0285	2314.5	85	0.0272	5657.0
Wbc	9	0.0212	7	0.0197	590.98	7	0.0163	573.76
Wbcd	31	0.0176	20	0.0114	744.43	18	0.0099	766.78

6.2 Results and discussion

In this section we present the experimental results obtained for the different approaches on datasets. Population size in clonal selection and firefly algorithms is equal to the features number in the database.

We can notice in Table 2 that solutions describe the datasets using both AIS and FFA are well. They produce a very good error rate using artificial neural network classifier with less number of features. We also compared the time costs of the feature selection methods, when the number of all available features is *n*, the time complexity of AIS generally corresponds to a large amount of time when *n* is big. The performance of FFA is better when the characteristics are increasing. When taking 166 features (Musk1 dataset), the error rate value and the number of select features are well.

6.3 Proposed approach

Table 3 presents the results of the proposed hybrid feature selection approach in terms of number of selected features and classification error for the ten databases.

Table 3 shows the experimental results using our hybrid algorithms on different datasets. We test several features and different values of parameters for the IFA 1 and IFA 2. The results demonstrate that the features selected by our IFA can accomplish the goal of achieving higher accuracy with smaller size or equal of features. Results show that the proposed methods are able to produce good performance on reducing the effects of the outliers and the noises and improve classification accuracy.

We can notice that solutions describe the datasets using both traditional (AIS, FFA) and hybrid approaches (IFA 1, IFA 2) are quite good. The hybrid approaches produce significant results both in terms of reduction of the selected features number and improvement of classification performance on the used databases.

The hybrid approaches have achieved the best results in terms of classification performance with the majority of databases.

Table 3. Results for Proposed algorithms

Data set	IFA 1		IFA 2	
	Number of features	Error	Number of features	Error
Spect Heart (Binary)	14	0.1122	11	0.0995
Spect Heart	28	0.0561	24	0.0564
Parkinson's Disease	10	0.0236	12	0.0298
Parkinson 2	13	0.1718	16	0.1686
Glass dataset	7	0.5100	8	0.3968
Breast tissue	7	0.3258	5	0.3151
Ionosphere	21	0.0267	22	0.0325
Musk1	94	0.0301	83	0.0243
Wbc	7	0.0179	6	0.0173
Wbcd	18	0.0086	21	0.0083

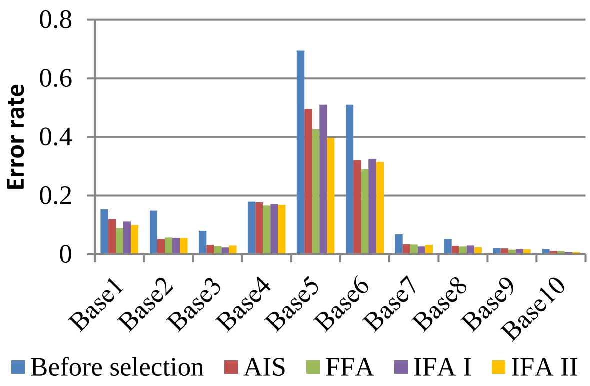

Figure.5 Error rates with algorithms used

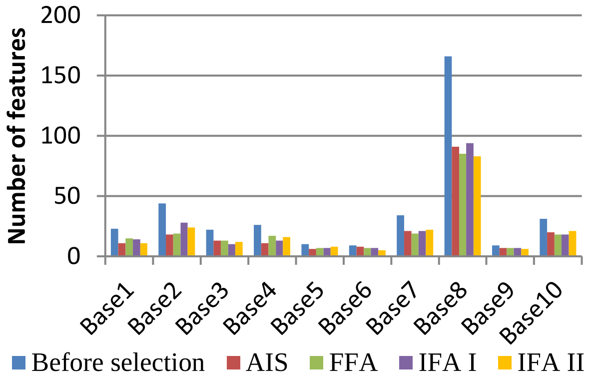

Figure.6 Number of selected features with algorithms used

The results of proposed algorithms show that the initial population influences the robustness and the convergence of AIS algorithm. In IFA 2 algorithm the initial population used is the generated using local search technique (FFA) instead a random generation. This step ensures diversity of the population, while the AIS tries to converge the initial solutions toward the optimal solution.

Also, the results of FFA 1 showed that the high exploratory mechanism of AIS and accurate local search of FFA combined are able to provide superior results when using to select relevant features.

7. Comparative study

In this section, the results of the proposed methods are verified against some of the methods in the literature.

Table 4 depicts the error classification rate and number of selected features of the proposed system when comparing with the existing methods.

The results showed the superior performance of the proposed approaches as compared to the existing algorithms in the literature.

The IFA algorithms benefit from high exploration due to the position updating equations of FFA. The solutions constantly face random changes using multiple best solutions. The search in feature selection problems changes for every dataset. IFA has a best performance in high dimension (Musk1) and large datasets (wbc and wbcd), and feature selection has a very significant result compared to the other methods.

Table 4. Previous works

Algorithms	Glass Dataset		Musk1		wbc		wbcd	
	Error	Number of features	Error	Number of features	Error	Number of features	Error	Number of features
IFA 1	0.5100	7	0.0301	94	0.0179	7	0.0086	18
IFA 2	0.3968	8	0.0243	83	0.0173	6	0.0083	21
ACO-ABC-FA/GA [32]	0.3922	8	0.0296	84	0.0165	8	0.0082	20
PSO [31]			0.2277	26	0.3483	2		
ACO-PSO1 [31]			0.2474	52	0.2550	2		
ACO-PSO2 [31]			0.2309	28	0.3033	2		
ACO-PSO3 [31]			0.2527	41	0.3033	2		
ALO [21]					0.031		0.028	
HBALO [21]					0.056		0.037	
BPSOWFSS [4]					0.2315	9		
Fast BPSOWFSS[4]					0.2308	9		

8. Conclusion

In this work a system to evaluate Feature Selection Algorithms was proposed in order to understand their general behavior on the particularities of relevance, irrelevance, redundancy and sample size of synthetic data sets. Thereafter, new methods for feature selection based on artificial immune system and firefly algorithms were proposed. The proposed algorithms were tested over 10 well-known UCI datasets and compared among commonly used feature selection methods: PSO, ACO, GA, ALO and their hybrid approaches. Simulation results show the superior performance of the proposed algorithms over the other algorithms. The error rate is obviously best than traditional feature selection approaches. As a whole, IFA has a stable performance in high dimension and large datasets, and feature selection has a very significant result. As future activities, this work can be extended in many ways to carry up richer evaluations and the use of combined evaluation measures.

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] M. Hall, "Feature Selection for Discrete and Numeric Class Machine Learning", In: Proc. of the *7th International conference on Machine Learning*, pp.359-366, 2000

[4] X. Liu and L. Shang, "A Fast Wrapper Feature Subset Selection Method Based on Binary Particle Swarm Optimization", *IEEE Congress on Evolutionary Computation*, pp.3347-3353, 2013.

[5] M. Mafarja and S. Abdullah, "Investigating memetic algorithm in solving rough set attribute reduction", *International Journal of Computer Applications in Technology*, Vol. 48, pp. 195-202, 2013.

[6] M. Anusha and J. Sathiaseelan, "Feature Selection Using K-Means Genetic Algorithm for Multi-objective Optimization", *Procedia Computer Science*, Vol. 57, pp. 1074-1080, 2015.

[7] 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.

[8] A. Al-ani, "Ant Colony Optimization for Feature Subset Selection", *World Academy of Science, Engineering and Technology,* Vol.1, No. 4, 2007.

[9] R. N. Khushaba, A. Al-Ani, A. Al-Sukker, and A. Al-Jumaily, "A combined ant colony and differential evolution feature selection algorithm", *Ant Colony Optimization and Swarm Intelligence*, pp.1–12, 2008.

[10] A. Unler and A. Murat, "A discrete particle swarm optimization method for feature selection in binary classification problems", *European Journal of Operational Research,* Vol. 206, No. 3, pp. 528–539, 2010.

[11] C. Şahín and B. Dírí, "Robust Feature Selection With LSTM Recurrent Neural Networks for Artificial Immune Recognition System", *IEEE Access,* Vol.7, pp. 24165 – 24178, 2019.

[12] G. Dudek, "Artificial Immune System With Local Feature Selection for Short-Term Load Forecasting", *IEEE Transactions on Evolutionary Computation,* Vol.21, No.1, pp.116-130, 2016.

[13] K. Verma, P. Srivastava, and A. Jain, "Clonal selection based AIS weighted feature extraction algorithm to identify the multiclass web pages categories", *International Journal of Innovative Technology and Exploring Engineering*, Vol.8, pp. 156 - 161, 2019.

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

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

[16] M. Mafarja and M. Mirjalili, "Whale optimization approaches for wrapper feature selection". *Applied Soft Computing,* Vol.62, pp.441–453, 2018.

[17] S. Arora and P. Anand, "Binary butterfly optimization approaches for feature selection", *Expert Systems with Applications*, Vol.116, 2018.

[18] 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.

[19] S. Mukherjee and L. Bhaumik, “Simultaneous Clustering and Feature Selection Using Nature-Inspired Algorithm”, *Lecture Notes in Networks and Systems*, pp. 545 – 550, 2019.

[20] 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

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

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

[23] H. M. Zawbaa, E. Emary, C. Grosan, and V. Snasel, “Largedimensionality small-instance set feature selection: a hybrid bioinspired heuristic approach”, *Swarm and Evolutionary Computation*, Vol. 42, pp. 29–42, 2018.

[24] L. Ladha, “Feature selection methods and Algorithms”, *International Journal on Computer Science and Engineering*, Vol.3, pp.1787 – 1797, 2011.

[25] I. Fistera, “A comprehensive review of firefly algorithms”, *Swarm and Evolutionary Computation*, Vol.13, pp.43-86, 2013.

[26] C. Blum and X. Li, *Swarm intelligence in optimization: Introduction and Applications*, pp.43-86, Springer Verlag, Berlin, 2008.

[27] H. Banati and M. Bajaj, “Fire Fly Based Feature Selection Approach”, *International Journal of Computer Science*, Vol. 8, No. 2, 2011.

[28] S. Dash, R. Thulasiram, and P. Thulasiraman, “Modified Firefly Algorithm With Chaos Theory for Feature Selection: A Predictive Model for Medical Data”, *International Journal of Swarm Intelligence Research*, Vol.10, No.2, pp.1-20, 2019.

[29] G. Somasundaram and M. Sabrigiriraj, “A hybrid multi-objective firefly and simulated annealing based algorithm for big data classification”, *Concurrency and Computation: Practice and Experience*, Vol. 30, 2018.

[30] L. N. de Castro and J. Timmis, “Artificial Immune Systems: A Novel Paradigm to Pattern Recognition”, In: *Proc. of Artificial Neural Networks in Pattern Recognition*, pp. 67-84, 2002.

[31] K. Menghour and L. Meslati, “Hybrid ACO-PSO Based Approaches for Feature Selection”, *International Journal of Intelligent Engineering and Systems*, Vol.9, No.3, pp.65-79, 2016.

[32] 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.