

A Combination Approach of Two Metaheuristic Algorithm for Optimal Feature Selection: Case Study Email Spam Detection

Hekmat Mohmmadzadeh

Department of Computer Engineering, Urmia Branch, Islamic Azad University, Urmia, Iran
Corresponding Author * {hekmoa@gmail.com}

Abstract

Selecting a feature in data mining is one of the most challenging and important activities in pattern recognition. The issue of feature selection is to find the most important subset of the main features in a specific domain, the main purpose of which is to remove additional or unrelated features and ultimately improve the accuracy of the categorization algorithms. As a result, the issue of feature selection can be considered as an optimization problem and to solve it, meta-innovative algorithms can be used. In this paper, a new hybrid model with a combination of whale optimization algorithms and flower pollination algorithms is presented to address the problem of feature selection based on the concept of opposition-based learning. In the proposed method, we tried to solve the problem of optimization of feature selection by using natural processes of whale optimization and flower pollination algorithms, and on the other hand, we used opposition-based learning method to ensure the convergence speed and accuracy of the proposed algorithm. In fact, in the proposed method, the whale optimization algorithm uses the bait siege process, bubble attack method and bait search, creates solutions in its search space and tries to improve the solutions to the feature selection problem, and along with this algorithm, Flower pollination algorithm with two national and local search processes improves the solution of the problem selection feature in contrasting solutions with the whale optimization algorithm. In fact, we used both search space solutions and contrasting search space solutions, all possible solutions to the feature selection problem. To evaluate the performance of the proposed algorithm, experiments are performed in two stages. In the first phase, experiments were performed on 10 sets of data selection features from the UCI data repository. In the second step, we tried to test the performance of the proposed algorithm by detecting spam emails. The results obtained from the first step show that the proposed algorithm, by running on 10 UCI data sets, has been able to be more successful in terms of average selection size and classification accuracy than other basic meta-heuristic algorithms. Also, the results obtained from the second step show that the proposed algorithm has been able to perform spam emails more accurately than other similar algorithms in terms of accuracy by detecting spam emails.

Keywords: *Classification, Hybrid, Whale Optimization Algorithm, Email Spam.*

1. Introduction

Feature selection can be defined as the process of identifying related traits and eliminating unrelated and repetitive traits with the aim of observing a subset of traits that describe the problem well and with minimal reduction in efficiency. This has many benefits, some of which are described below. To define a "Feature Relevance" problem, the definition is used: "A attribute is relevant if it contains information about the target." More formally, John and Kohavi have divided the traits into three distinct categories: "strongly relevant," "weakly relevant," and "unrelated." These are called irrelevant features. In the approach of these researchers, the relationship between the X feature is defined as an ideal bisexual classification. Characteristic X is strongly related when its removal leads to damage to the accuracy of the prediction of the ideal biodegradation. This feature is called weakly related if it is not strongly related and there is a subset of S-properties, so that the performance classification of the ideal ideal on S is worse than the efficiency of $\{S \cup \{X\}$. An unrelated feature is defined if it is not strongly and poorly related. One feature is usually considered to be Feature Redundancy if there is a correlation between the features. The notion

that two features overlap is acceptable to many researchers if their values are completely correlated, but at the same time it may not be easy to recognize the excess of features when a feature is associated with a set of features. According to the definition provided by John and Kohawi, a property is redundant if it is to be removed, so that it is weakly related and has a Markov blanket within the set of current properties. Because unrelated features need to be removed from all sides, they are cleared according to this definition. The existence of thousands of programs of information systems has complicated the role of extracting useful information from the collected data [1]. Feature selection (FS) is one of the most important steps in pre-processing because its purpose is to eliminate redundant and irrelevant variables in a data set. Feature selection methods are classified as packaging and filter [2, 3]. A number of approaches proposed for FS can be broadly classified into the following three categories: packaging, filter, and hybrid. In the packing method, a predetermined learning model is assumed, in which features are selected that justify the learning performance of a particular learning model, while in the filtering method, statistical analysis of the set of characteristics, no learning model is required. Schematic diagrams show how to find packing approaches and filter outstanding features in Figure 1. The combined approach tries to take advantage of the complementary strengths of packaging and filtering approaches [6, 7].

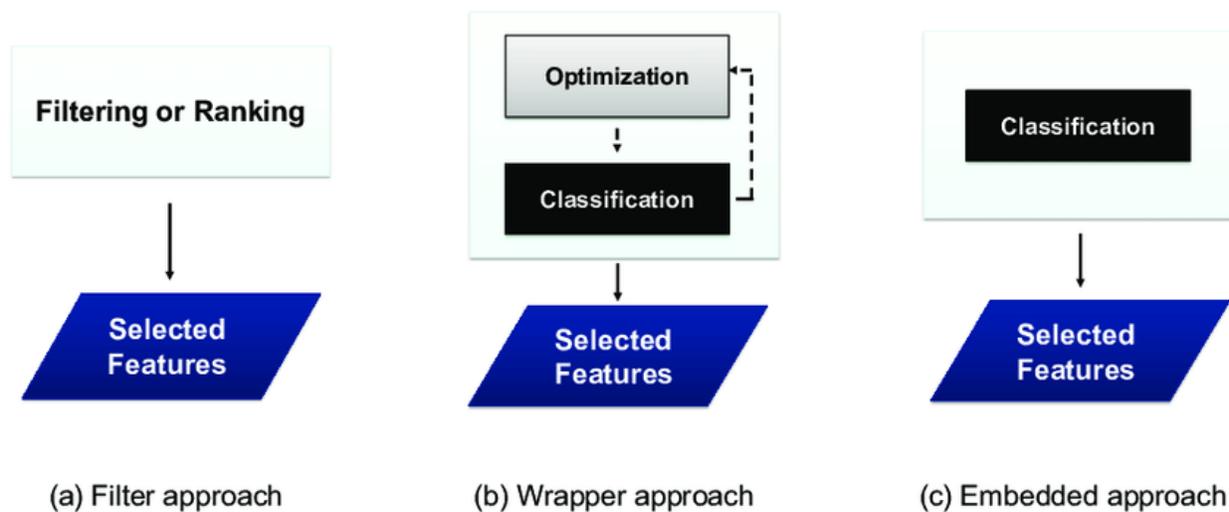


Fig. 1. Schematic diagrams of (a) wrapper approach and (b) filter approach [7].

Classical optimization techniques have limitations to solve feature selection problems [8] and, therefore, evolutionary computational algorithms (EC) are alternatives to solving these limitations and finding the best solution [9]. Evolutionary computational (EC) algorithms are inspired by nature, group dynamics, social behaviors, and species biological interaction in a group. The binary version of these algorithms allows us to examine problems such as feature selection and achieve superior results. Thanks to feature selection techniques, we have improved benefits such as better model interpretation, shorter training times, and generalization by reducing additional connections when building classification models. FS can be considered as a space search. Therefore, a full search can be performed in all covered search spaces. However, this approach is not possible for a large number of features. Therefore, an exploratory search evaluates these features that have not yet been selected in each iteration. A random search creates random subsets in the search space that can be evaluated to determine the importance of rank performance [10]. Nature-inspired meta-heuristic algorithms are now one of the most widely used algorithms for solving optimization problems. Optimization is the process of finding optimal solutions to a particular problem. There are many nature-inspired metaheuristic algorithms that work on the feature selection problem, some of which are combined to combine algorithms, and some are used alone to solve this problem. Due to their random nature, metaphors such as particle optimization (PSO) [11], evolutionary algorithms (EA) [12], milk optimization (ALO) [13], bat algorithm (BA) [14],

Flame worm algorithm (FA) [15], Whale Optimization Algorithm (WOA) [16], Genetic Algorithm [17], Flow Pollution Algorithm (FPA) [18] and other metaphorical algorithms are used. Characteristics Selection Several previous studies have used hybrid algorithms for the problem of feature selection. In Section 2, we will describe hybrid algorithms for combining features.

The rest of this document is organized as follows: Section 2 provides related work. The basic concepts of the WOA and FPA algorithms are presented in Section 3. Section 4 details the proposed method. In Section 5, the experimental results are presented and the results are analyzed. Finally, in Section 6, conclusions and subsequent work are given.

2. Related Works

In recent years, the metaheuristic hybrid has been used by many researchers in the field of optimization. In the feature selection domain, many successful hybrid metaheuristic algorithms have been proposed [2, 7, 10, 19-21].

The first hybrid method for the selection of characteristics belongs to Oh et al. [22] incorporating local search operations that are Forward Sequential Search (SFS), Floating Forward Sequential Search (SFFS), and Polynomial Time Approach (PTA) in GA to fine-tune the search process. Hybrid GA shows better convergence property over standard GA in experiments performed on various UCI repository data sets, including Glass, Vowel, Wine, Letter, Vehicle, Segmentation, WDBC, Ionosphere, Satellite, and Sonar. Khushaba et al. [23] combined ACO and DE for feature selection, where DE was used to search for the optimal feature subset based on the solutions obtained by ACO. Olabiyisi et al [24] develop another novel hybrid algorithm that includes GA and SA metaheuristic to extract characteristics for timing problems. In the proposed algorithm, the SA selection process is used instead of the GA selection process to avoid a local optimum. Experimental results show that SA performs better than GA and GA-SA hybrid in terms of optimization and runtime; however, GA and SA execution times are higher than those of the hybrid method. Due to the runtime performance, it is concluded that the hybrid algorithm is more applicable compared to GA and SA. In [25], a filter measurement (Pearson correlation measure) and an envelope measure (classification precision) were combined to form a single fitness function in an AG for the selection of characteristics to take advantage of each measurement. Akila [26] constructs a hybrid envelope and filter characteristics selection algorithm for classification problems using a combination of GA and local search (LS) technique. In the hybrid method, first, LS is applied by using correlation-based filter methods that include discretization, classification, and redundancy elimination with a measure of symmetric uncertainty for subsets of features; and then standard GA operators are applied to these subsets. Experimental analysis performed on the DNA gene analysis data set obtained from the UCI deposit shows that the hybrid method has the best performance. Babatunde et al. [27] propose another hybrid algorithm based on ACO and GA for the selection of characteristics. In this new algorithm, selected entity subsets are evaluated using the SVM classifier. The proposed algorithm is tested on the face detection dataset and compared to standard ACO and GA. The results demonstrate that the new hybrid technique outperforms ACO and GA.

Hasani et al. [28] developed a combination of Linear Genetic Programming (LPG) and Bee Algorithm (BA) for feature selection in Intrusion Detection Systems (IDS). In the proposed algorithm, LPG is used to generate solutions of feature subsets, then the BA neighborhood search process is applied to these solutions; and finally, SVM is used to evaluate subsets of features. Experimental results demonstrate that the LPGBA method increases the accuracy of the classification and is more efficient than the basic LPG and BA. In [29] Liao et. to the. It proposes a new feature selection algorithm based on Ant Colony Optimization (ACO), called Advanced Binary ACO (ABACO), and it is introduced. The simulation results verify that the algorithm provides a suitable subset of features with good classification accuracy using a smaller feature set than the competitive feature selection methods.

Nekkaa and Boughaci [30] propose a new hybrid search method combining the harmony search algorithm (HSA) and stochastic local search (SLS) for the selection of characteristics in the classification task. A new probabilistic selection strategy is used to apply stochastic exploitation. This algorithm is wrapped with the SVM classifier. Experimental results show that the HSA-SLS method is better than HSA and GA for feature selection. Sayed et al. [18] the study has investigated the effectiveness of combining CSA with FPA to compose a new hybrid algorithm called Binary

Clonal Pollination Algorithm of Flowers (BCFA) to solve the problem of selection of characteristics. The precision of the OPF classifier was used as an objective function to maximize.

The experiments were performed on three mass data sets in the UCI machine learning tank. The results showed that BCFA can achieve the best classification quality using the minimum number of properties selected in a short period of time. Khoshaba et al. [5] A new hybrid approach has been proposed that combines the artificial colony optimization method of artificial bees with a differential evolutionary algorithm for selecting classification tasks. The developed hybrid method is evaluated using fifteen sets of data from the UCI repository, which is widely used in classification problems. The experimental results of this study show that our developed combined method is able to select good features for classification works to improve the working time and accuracy of the classifier.

In [2] Mafarja et al. Suggest the use of two separate hybrid models in the development of different feature selection methods based on the KIT optimization algorithm (WOA). The effectiveness of the proposed methods in a database of 18 standard criteria in the UCI warehouse is evaluated and compared with the methods of selecting the three known packaging features in the literature. Experimental results confirm the effectiveness of the proposed methods in increasing the accuracy of classification compared to other packaging algorithms. Hassani et al. [21] Introduces a new blending algorithm based on the new Swarm AC-ABC, which uses the ant colony algorithms of ants and artificial bee colonies (ABC) to optimize the selection of features. Thirteen UCI standard databases (University of California, Irvine) were used to evaluate the proposed algorithm. Experimental results show promising behaviors in increasing the accuracy of classification of the proposed method and in the optimal selection of features. To increase the accuracy of the classification, Liao et al. Creates five discrete search algorithms (SOS) search algorithms to optimize the properties and adjacent dimensions of the neighboring k-close model at the same time. Using the adjacent K-close classification model, the performance of these algorithms in terms of classification error and calculation time is evaluated based on the validation of the K-class classified cross using 11 data sets [19].

Hassani et al. [1] Spatial optimization methods for the dynamics of the existing evolutionary population and the problems of feature selection. Suggested methods are used to solve a variety of selection tasks. The proposed approaches are compared in the UCI 22 data set. Comparative experiments show the superiority of the proposed approaches over other similar methods in the literature. Khoshaba et al. [20] To solve the feature selection problems, introduce a new binary combination of bats and an algorithm corresponding to advanced particles. The proposed algorithm has been compared with the main optimizers and other optimizers used in the selection of features in the past. A set of evaluation indicators are used to evaluate and compare different optimizers in 20 standard data sets from the UCI repository. The results show the ability of the proposed algorithm to search for field properties to combine optimal features.

3. Material and Methods

Optimal whale and flower pollination algorithms are two examples of meta-heuristic algorithms that have been used in combination to select features based on case-based learning and case studies on spam email detection. In the following discussion, we will describe and examine each of these algorithms.

3.1. Whale Optimization Algorithm

The Whale Optimization Algorithm (WOA) is a search and optimization algorithm recently developed by [31]. This is a mathematical model of the movements and behaviors of small whales in search of food. The performance of the WOA algorithm is inspired by an attack strategy in which whales are trapped in bubbles, they start targeting fish and form spiral bubbles around the fish, 12 meters below the ground. They landed and then retreated to catch themselves. Boat Kurdish fish as shown in Figure 2. The search process in this algorithm is described by random search according to the relative positions of the whales. Instead of choosing the best option, you can mathematically translate old solutions mathematically through another random selection. In addition to these interesting behaviors, WOA is very different from other optimization algorithms and requires only two parameter settings. These parameters make it possible to continuously transfer between search and operation processes.

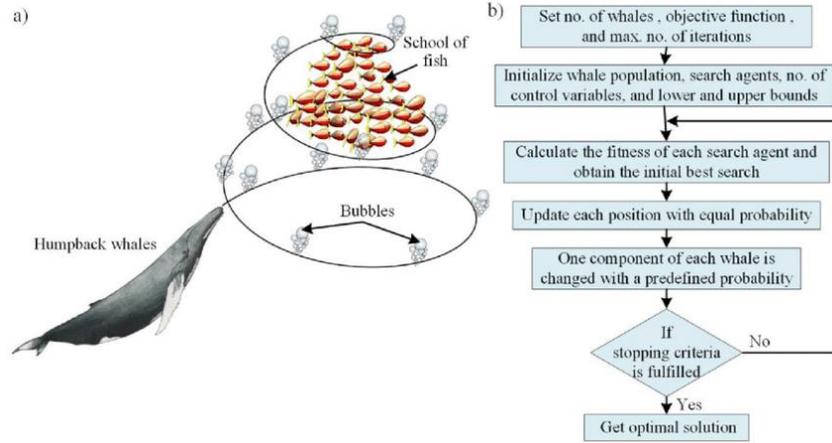


Fig.2. Bubble-net feeding behavior of humpback whales [31].

The mathematical model of encircling prey, spiral bubble-net foraging maneuver and search for prey is described in the following section:

3.1.1. Encircling prey

Humpback whales encircle the prey and update their position towards the best search agent with the increasing number of iterations from start to a maximum number of iterations. This behavior is mathematically formulated as:

$$\vec{D}_i = |\vec{C} \cdot \vec{X}^*(t) - \vec{X}(t)| \quad (1)$$

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \quad (2)$$

Where \vec{A} and \vec{C} are the coefficient vectors, t indicates the current iteration, X^* is the position vector of the best solution obtained so far, \vec{X} is the position vector, $| |$ is the absolute value and \cdot is an element-by-element multiplication. The vectors \vec{A} and \vec{C} are calculated as follows:

$$\vec{A} = 2\vec{a} \cdot \vec{r} - \vec{a} \quad (3)$$

$$\vec{C} = 2 \cdot \vec{r} \quad (4)$$

3.1.2. Bubble-net attacking method

The following two methods have been developed for mathematical modeling of pure bubble whale movement:

1. Decreased environmental mechanism: This behavior is achieved by lowering the E value from 2 to 0. (3) During repetition. The new location of the search agency can be determined anywhere between the original state of the agent and the location of the best agent by assigning random values to $[-1, 1]$.

2. Spiral repositioning position: To simulate whale movement, the spiral equation between prey and whale position is as follows:

$$\vec{X}(t+1) = \vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) \quad (5)$$

The pseudo code of the WOA algorithm is illustrated in Fig 3.

```

Randomly initialize the whale population.
Evaluate the fitness values of whales and find out the best search agent  $X^*$ .
while  $t < t_{\max}$ 
    Calculate the value of  $a$  according to Equation (13)
    for each search agent
        if  $h < 0.5$  then
            if  $|A| < 1$  then  $X(t+1) = X^*(t) - A \cdot D$ 
            else if  $|A| \geq 1$  then  $X(t+1) = X_{\text{rand}}(t) - A \cdot D''$ 
            end if
        else if  $h \geq 0.5$  then
             $X(t+1) = D' \cdot e^{bl} \cdot \cos(2\pi l) + X^*(t)$ 
        end if
    end for
    Evaluate the fitness of  $X(t+1)$  and update  $X^*$ 
end while

```

Fig. 3. Pseudo-code of the WOA algorithm.

It should be noted that scrubber whales swim in a shorter frame around the prey and along the spiral path. Therefore, we estimate a 50% probability of choosing a short rotation method or a spiral model to update the position of the whales to model this behavior. The mathematical model is as follows:

$$\begin{aligned} \vec{X}(t+1) &= \{\vec{X}^*(t) - \vec{A} \cdot \vec{D}\} & \text{if } p \leq 0.5 \\ \{\vec{D}' \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(v)\} & & \text{if } p \geq 0.5 \end{aligned} \quad (6)$$

where $D' = |\vec{X}^* - \vec{X}(t)|$ and indicates the distance of the i th whale to the prey (the best solution obtained so far), b is constant for defining the shape of the logarithmic spiral, l is a random number in $[-1, 1]$ and represents a random number in $[0, 1]$.

3.1.3. Search for prey (exploration phase)

The variation of \vec{A} vector can be utilized to search for prey, i.e., exploration phase. Therefore, \vec{A} can be used with the random values greater than 1 or less than -1 to force search agents to move away from a reference whale. The mathematical model for this phase is as follows:

$$\vec{D} = |\vec{C} \vec{X}_{\text{rand}} - \vec{X}| \quad (7)$$

$$\vec{X}(t+1) = \vec{X}_{\text{rand}} - \vec{A} \cdot \vec{D} \quad (8)$$

Where \vec{X}_{rand} is a random position vector (a random whale) chosen from the current population .

3.2. Flower Pollination Algorithm

The Flower Pollination Algorithm was introduced in 2012 by computer scientist Zing Xi Yang, who has developed various algorithms such as Shabbat Cream, Bats and Coco.

Pollination of flowers is a process that involves the transfer of pollen from flowers. The main participants in such migration are birds, bats, insects and other animals. There are a number of flowers and insects that we can call pollination. These flowers can only attract birds that participate in this partnership, and these insects are the main pollinators of these flowers [32]. There are two types of pollination: antibiotic and non-live. Biotic flowers make up 90% of pollinators, while non-living plants make up 10%. Non-living pollination does not require pollination. Some insects visit certain types of flowers. However, these insects move away from other flower species and this phenomenon is called: flower stagnation [33]. All flowers that have the properties of perennial flowers guarantee maximum reproduction. This process is useful for both flowering and pollination. The pollinator receives enough nectar, while the pollen retains the pollen and thus increases the number of known flower species. Depending on the

availability of pollinators, pollination can also be defined as spontaneous and reciprocal pollination. In self-pollination, there is no reliable pollination, and during cross-pollination, pollutants such as bees, birds, and bats fly long distances, causing global pollination [34]. In [35], Young simulates the process of biological pollination using the following four laws of idealization:

- A) Environmental pollination and pollination is a global pollination with a pollinator that performs Levi's flight.
- B) Living and self-pollination are considered local pollination.
- C) The persistence of flowers can be considered as a possibility of reproduction. It is suitable for two flowers.
- D) Local pollination and global pollination are managed with the possibility of P [0, 1] transmission. Due to other factors such as physical proximity and wind, local pollination can play an important role in overall pollination activities.

To shape up-to-date formulas, we need to translate the above rules into updated equations. For example, in the global pollination stage, pollen gametes are carried by pollinators such as insects, and pollen can travel long distances because insects can often fly and travel much longer distances. Therefore, Rule 1 and the stability of flowers can be expressed mathematically as follows:

$$x_i^{(t+1)} = x_i^t + \gamma L(\lambda)(x_i^t - B) \quad (9)$$

Where x_i^t is pollen i or solution vector X_i at iteration t , and B is the current best solution found among all solutions at the current generation/iteration. Here γ is a scaling factor to control the step size. In addition, $L(\lambda)$ is the parameter that corresponds to the strength of the pollination, which essentially is also the step size. Since insects may move over a long distance with various distance steps, we can use a Lévy flight to imitate this characteristic efficiently. That is, we draw $L > 0$ from a Levy distribution:

$$L = \frac{\lambda \cdot \Gamma(\lambda) \cdot \text{Sin}(\lambda)}{\pi} \frac{1}{S^{1+\lambda}} (S \geq S_0 \geq 0) \quad (10)$$

Here, $\Gamma(\lambda)$ is the standard gamma function, and this distribution is valid for large steps $s > 0$. Then, to model the local pollination, both Rule 2 and Rule 3 can be represented as:

$$x_i^{(t+1)} = x_i^t + U(x_j^t - x_k^t) \quad (11)$$

Here from and from are pollinated from different flowers of similar plant species. This is actually an imitation of the permanence of flowers in a limited neighborhood. If mathematically, if x_j^t and x_k^t are of the same species or are selected from the same population, we obtain U from the equal distribution [0, 1], which can be a local random walk. Although flower pollination measures are carried out on a global and global scale, flower patches or adjacent flowers near local pollen are pollinated farther away. To mimic this, we can effectively use the law of 4 or p to use global public pollination to severe local pollination such as the possibility of proximity. To get started, we can use the simple value $p = 0.5$ as the initial value. The initial parameter $0.8 = p$ showed that it could perform better for most applications [35]. Figure 4 shows the fake code of the flower pollination algorithm.

Algorithm: *Flower Pollination Algorithm*

Require: *Set of population of n flowers/pollen gametes with random solutions*

Find the best solution g^ in the initial population*

Ensure: *Define a switch probability $P = 0.5 \frac{(Maxiteration - t)}{Maxiteration}$*

```

1:  Input: PM list, VM, set of parameters
2:  Output: VM allocation
3:  Execute: Objective Max  $R^A = (RU^{DC}) (DC_U^{Energy})$ 
4:  Initialize a population of n flowers/pollen gametes with random solutions
5:  Find the best solution  $g^*$  in the initial population
6:  Define a switch probability P
7:  While ( $t < MaxGeneration$ )
8:      For  $i = 1:n$  (all n flowers in the population)
9:          If  $rand < p$ ,
10:             Draw a (d - dimensional) step vector L which obeys a Levy distribution
11:             Global pollination via  $X_i^t = X_i^t + L (g^* - X_i^t)$ 
12:             else
13:                 Draw  $\epsilon$  from a uniform distribution in [0, 1]
14:                 Randomly choose j and k among all the solutions
15:                 Do local pollination via  $X_{i+1,G} = X_{i,G} + \emptyset (X_{n-best_iG}) + \vartheta (X_{p,G} - X_{q,G})$ 
16:                 where  $\emptyset = \vartheta = \epsilon$ ;
17:             end if
18:             Evaluate new solutions
19:             If new solutions are better, Then
20:                 Update them in the population
21:             end for
22:             Find the current best solution  $g^*$ 
23: End while

```

Fig. 4. Pseudo-Code Flower Pollination Algorithm.

The secondary flower pollination algorithm is provided by [36], in which the search area is designed as a two-dimensional bovine network in which solutions are updated in the corners of the mushroom cloud. The problem with selecting properties is whether or not you select a particular attribute, so the solution is presented in the form of a binary vector, where 1 indicates the property selected to create a new data set, and so on. The sigmoid function is used to construct this binary vector using the following equation:

$$s(x_i^j(t)) = \frac{1}{1 + e^{-x_i^j(t)}} \quad (12)$$

Thus, Eq. (9) and Eq. (11) will be replaced by the following equation:

$$x_i^j(t) = \begin{cases} 1 & \text{if } s(x_i^j(t)) > \sigma \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

In which $x_i^j(t)$ represents the new pollen (solution) i with the j^{th} feature vector, where $i = 1, 2 \dots m$ and $j = 1, 2 \dots d$, at the iteration t and $\sigma \sim U(0, 1)$. Fig.5 presents the pseudo-code of Binary Flower Pollination Algorithm.

```

1: Objective min or max  $f(x)$ ,  $x = (x_1, \dots, x_n)$ ;
2: Initialize a population of  $n$  flowers/pollen gametes with random solutions;
3: Find the best solution  $g^*$  in the initial population;
4: Define a switch probability  $p \in [0, 1]$ ;
5: while ( $t < MaxGeneration$ ) do
6:   for  $i = 1 : n$  (all  $n$  flowers in the population) do
7:     if  $rand < p$ , then
8:       Draw a ( $d$ -dimensional) step vector  $L$  which obeys a Lévy flight distribution;
9:       Global pollination via Eq. (1);
10:    else
11:      Draw  $\epsilon$  from a uniform distribution in  $[0, 1]$ ;
12:      Randomly choose  $j$  and  $k$  among all the solutions;
13:      Do local pollination via Eq. (3);
14:    end if
15:    Evaluate new solutions;
16:    if new solutions are better, update them in the population;
17:  end for
18:  Find the current best solution  $g^*$ ;
19: end while

```

Fig. 5. Pseudo-code of Binary Flower Pollination Algorithm [36].

3.3. Opposition Based Learning

The first concept of opposition was first expressed in ancient Chinese philosophy in the symbol of Yin Yang (Figure 6). This symbolizes the concept of a pair of black and white yin (receiving power, subtle, dark, passive) and yang (creative, masculine, light and active power). He also described the classical Greek elements of the patterns of nature (Figure 7) as contradictory concepts such as fire (hot and dry) and water (cold and wet), earth (cold and dry) and air (hot and humid). The nature of cold, hot, humid and dry and opposites shows them [37].



Fig. 6. Early opposite concept was mentioned in the Yin-Yang symbol [37]

Opposition-based learning (OBL) [38] is an effective concept for strengthening various optimization methods. The main idea of the OBL is to simultaneously consider the relevant antitrust assessment as secondary candidate solutions to achieve a better approach to the candidate's current solution. It has been proven that the solution of the opposing candidate randomly increases the chance of approaching a universally acceptable solution instead of a solution of the selected candidate. The opposite number can be identified as follows:

Def 1: Training based on the opposition of Rahnamayan et al. [39] This changes the existing search to a new search. If x a solution exists in a search field, where. $\in [L,.]$, Then yangi * in the modified space according to the opposite learning model, the new solution is as follows.

$$x^* = L + U - x \quad (14)$$

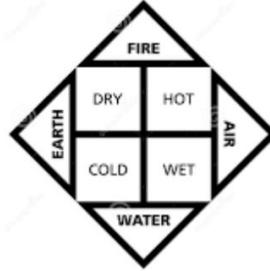


Fig. 7. The Greek classical elements to explain patterns in the nature [37]

4. Proposed Approach

In this section, the proposed method is described, which includes a combination of whale optimization algorithms and flower pollination. In the proposed method, we have used the binary version of the whale optimization algorithms [2] and flower pollination [36] to select the feature. In the proposed method, we also used the concept of opposition-based learning to contrast the search space and create better and newer solutions to improve the hybrid algorithm. In this section, we first describe the objective function selection function for the proposed algorithm and other meta-innovative algorithms in this paper. This is because feature selection can be considered as a multi-objective optimization problem in which two conflicting goals are achieved, which include the minimum number of selected features and higher category accuracy. Therefore, we need a categorization algorithm to define the objective function of the property selection problem, and since most researchers [1, 2, 10, 16, 19] use the simplest classification method, the KNN categorizer [40]. Benefiting, we also used this category to define the target function of the feature selection problem.

As a result, in the proposed method, we used the KNN classifier to more accurately evaluate the features selected by the proposed algorithm and other algorithms. Each solution is evaluated based on the proposed multi-objective function, which depends on the KNN categorizer. In the proposed multi-objective function, in order to balance between the number of features selected in each solution (minimum) and the accuracy of the classification (maximum), the fit function in Equation (15) is used to evaluate a solution in each meta-heuristic algorithm.

$$Fitness = \alpha \gamma_R(D) + \beta \frac{|R|}{|N|} \quad (15)$$

Where $\alpha \gamma_R(D)$ shows the classification error rate of a category and also $|R|$ The multi-line subsystem is selected and $|N|$ The total number of features in the data set is also the α parameter, the importance of the category quality, and the β parameter of the subset length. The values of these two parameters are also obtained as $\alpha \in [0, 1]$ and $\beta = (1 - \alpha)$ from the study [13].

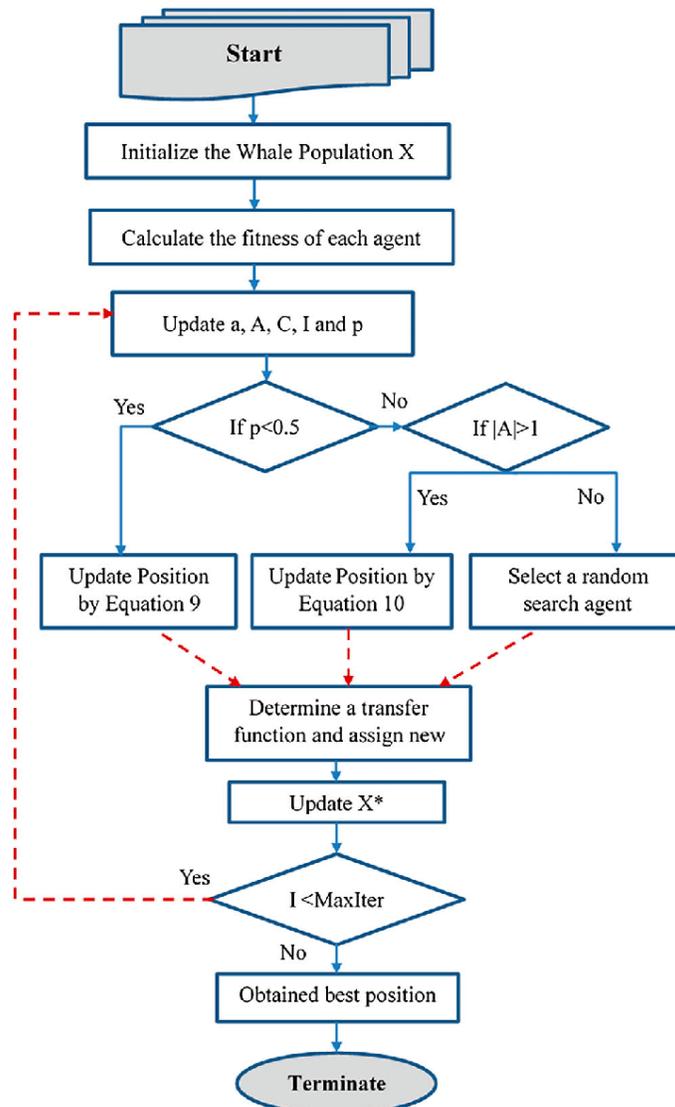


Fig. 8. Opposition-Based Hybrid Whale Optimization Algorithm.

After defining the objective function, we first used the combined method of whale optimization and flower pollination algorithms and then the proposed method of opposition-based learning concept to improve the hybrid algorithm, which we will discuss as the proposed method in this article. The main purpose of combining metaheuristic algorithms is to make more use of the natural process of two different algorithms to better solve all kinds of optimization and difficult problems. When combining meta-heuristic algorithms, the hybrid algorithm maintains the balance and performance of the hybrid algorithm by maintaining a balance between exploration and productivity. The whale optimization algorithm is a recently introduced optimization algorithm that shows very good results in solving many optimization problems. However, exploration in the standard WOA algorithm (Equation (8)) depends on changing the position of each search agent based on a randomly selected solution. It also gives us a glimpse into the productivity of the flower pollination algorithm to better improve the WOA algorithm in the feature selection problem.

Due to the fact that all non-innovative algorithms, including whale optimization and flower pollination algorithms, use random quantification or uniform distribution to create the initial population, sometimes the production of inappropriate primary population causes less innovative initiatives towards The optimal target is convergent, and many of the measured variables, such as computational time, storage memory, and complexity, are related to the distance of this initial value from the optimal global solution. If we can examine a solution at the same time and in contrast to that

solution, in fact, by using a population and the opposite of that population in meta-heuristic algorithms, great potential is created to accelerate convergence and improve accuracy in these algorithms. As a result, we have ensured the convergence and accuracy of the proposed algorithm by adding the concept of learning based on opposition to our combined algorithm. Figure 8 shows how the combination of whale optimization algorithms and flower pollination based on opposition-based learning is proposed.

As a result, in our proposed method, we tried to solve the problem of optimization of feature selection by using natural processes of whale optimization and flower pollination algorithms. We used opposition. In the proposed method, whale optimization algorithms use the bait siege process, bubble attack method and bait search, create solutions in their search space and try to improve it, and along with this algorithm, flower pollination algorithm by creating a way. Opposite solutions The whale optimization algorithm operates on the principle of opposition-based learning. In fact, in continuation of the flower pollination algorithm with two global and local search processes, the solutions created in the contrasting space will improve the solutions of the whale optimization algorithm. Also, at the end of each generation of the proposed algorithm, the population of the whale optimization algorithm population is combined with its opposite population, i.e. the population of the pollen pollination algorithm, and only solutions for the next generation are selected that are more optimal in terms of objective function. Thus, by combining two optimization algorithms and using contrasting search space, we were able to provide a powerful algorithm with high convergence and accuracy in this article. The simulation results in Section (5) confirm the proposed algorithm and convergence speed and accuracy of the proposed algorithm with various experiments.

5. The experiment results

The experiments were carried out using a PC Intel(R) Core(TM) i5-2430M CPU 2.40 GHz with 6GB RAM, Windows 10 operating system. The implementation of the proposed algorithm is done using Matlab.

To evaluate the performance of the proposed algorithm in this section, experiments are performed in two stages. In both cases, a binary version of the comparison algorithm is used. In the first phase, experiments were performed on 10 sets of data selection features from the UCI data repository

Be. In Section 5.1, we provide a complete overview of the first stage experiments, along with the name and features of the UCI data set, and the initial settings and results of the proposed algorithm on the UCI data set.

In the second step, we tried to test the performance of the proposed algorithm by detecting spam emails. In Section (5.2), the experiments related to the second stage, along with the name and characteristics of the spam email data set and the initial settings and the results of the proposed algorithm, are fully listed on the spam email data set. The results obtained from the first step show that the proposed algorithm, by running on 10 UCI data sets, has been able to perform more successfully than other algorithms in terms of average selection number of features and classification accuracy. Also, the results obtained from the second step show that the proposed algorithm has been able to perform spam emails more accurately than other similar algorithms in terms of accuracy.

5.1. UCI Datasets

To evaluate the effectiveness of the proposed approaches, experiments were performed on 10 FS standard databases in the UCI database [42]. To analyze the performance of the algorithm presented in terms of features and samples, a low-dimensional data set with low and high dimensions and subsequent and small and small in this document has been approved. The packaging method is based on the KNN classification (where $K = 5$ [13]) is used to produce the best reduction. In the proposed method, each data set is divided into a cross-comparison method for evaluation [43]. In the inspection of K-folding stocks, K-1 is used equally for drills and verification, the rest is used for folding testing. This M load process is repeated. The dimensions of the training set and the test sample are the same. Table 1 shows the details of the data set used, for example, the number of features and samples in each set. These include the chest, wine, bottle, contraceptive selection (CMC), system entry (heart), ionosphere, lymphography, spectrum, blood, zoo.

Table 1: List of Datasets Used in the Experiments.

NO	Dataset	No. of attributes	No. of objects
1.	Boold	4	748
2.	BreastEW	30	569
3.	CMC	9	1473
4.	Glass	10	214
5.	Heart	13	270
6.	Ionosphere	34	350
7.	lymphography	18	148
8.	Spect	22	267
9.	Wine	13	178
10.	Zoo	16	101

5.1.1. Comparison of WOA, FPA and HWOAFPA

The performance of WOA, FPA and HWOAFPA in terms of average selection criteria and category accuracy are discussed in this section. In the first experiment in this section, the parameters are set as follows. The maximum number of repetitions is 25 and the population size is 20. The results of WOA, FPA and HWOAFPA are shown in Table (2).

Table 2: Classification accuracy and average selected attributes obtained from in test1 with 25 iteration

Dataset	Accuracy			Attributes		
	WOA	FPA	HWOAFPA	WOA	FPA	HWOAFPA
Boold	0.7086	0.6845	0.7086	2.00	2.40	2.00
BreastEW	0.9439	0.9263	0.9474	13.7	19.2	13.5
CMC	0.4898	0.4681	0.4912	4.95	5.80	5.00
Glass	0.9907	0.9907	0.9907	2.00	6.30	1.85
Heart	0.8148	0.6222	0.8222	2.95	8.30	4.10
Ionosphere	0.9261	0.8523	0.9205	4.90	22.1	12.7
lymphography	0.8784	0.6757	0.9189	11.0	11.1	7.00
Spect	0.6791	0.5896	0.6716	2.00	14.9	6.80
Wine	0.9326	0.5843	0.9387	8.00	7.85	6.20
Zoo	0.9216	0.8824	0.9804	6.00	10.4	9.00

Examining Table 2, it is clear that the proposed algorithm performs better than the whale optimization and pollen algorithms of the flowers to achieve the two main objectives, which include the accuracy of the classification and the number of selected features. The proposed algorithm in this experiment is better among the 10 UCI data sets in terms of accuracy in 7 data sets than the whale optimization and flower pollination algorithms, and also the proposed algorithm is better in terms of the number of features selected in the 6 data sets. Optimize whale optimization algorithms and flower pollination. From the results reported in this experiment, according to Table (2), it can be concluded that the hybrid model, using the opposition-based learning method, has a significantly higher performance than the whale optimization algorithms and flower pollination. Because the opposition-based learning method in the proposed method is between repeating the whale optimization algorithms and pollinating the flowers, better results can be achieved by increasing the number of repetitions in the proposed algorithm. This motivated us to examine the

proposed approach with a maximum number of iterations of 50 and a population size of 10. The results of the proposed algorithm with a higher number of iterations are shown in Table (3).

Table 3: Classification accuracy and average selected attributes obtained from in test2 with 50 iteration

Dataset	Accuracy			Attributes		
	WOA	FPA	HWOAFPA	WOA	FPA	HWOAFPA
Boold	0.6604	0.6016	0.7086	3.00	2.60	2.00
BreastEW	0.9579	0.9404	0.9614	18.0	20.4	16.0
CMC	0.5020	0.4695	0.5115	3.20	5.10	6.40
Glass	0.9907	0.9813	0.9907	4.00	6.40	2.20
Heart	0.7704	0.5037	0.8593	2.00	8.90	4.00
Ionosphere	0.8977	0.8693	0.9148	7.00	22.3	13.0
lymphography	0.8378	0.7432	0.8649	9.00	11.5	9.00
Spect	0.7015	0.4627	0.7761	10.1	13.7	3.00
Wine	0.9438	0.7079	0.9551	10.0	8.80	3.00
Zoo	0.8824	0.6667	0.9608	9.00	10.8	12.0

Examining Table 3, it is clear that as the number of repetitions increases, the proposed algorithm of Whale Optimization and Flower Pollination Algorithms works much better to achieve the two main objectives, which include classification accuracy and the number of selected features. Slow down. In particular, this experiment shows that in order to accurately categorize, the proposed algorithm out of 10 UCI data sets, in all data sets, perform better than whale optimization and flower pollination algorithms, and also the proposed algorithm in terms of the number of selection features in 7 Data sets work better than whale optimization and flower pollination algorithms. From the results reported in this experiment, according to Table (3), it can be concluded that the hybrid model has greatly increased the performance of whale optimization algorithms and flower pollination in terms of convergence and accuracy by using opposition-based learning methods. Since the opposition-based learning method is located between the repetition of the whale optimization algorithms and the pollination of the flowers, it can be seen that by increasing the number of repetitions, better results have been obtained for the proposed algorithm.

From the results reported in the two experiments performed in this section, it can be concluded that the proposed algorithm has solved the problem of optimizing the selection of features well, and also in the last experiment, considering the success of the algorithm in all data sets, convergence speed and algorithm accuracy. The proposal is guaranteed by the use of opposition-based learning methods. In the next section, we will try to compare the proposed algorithm with other basic algorithmic algorithms for further evaluation.

5.1.2. Comparison with the State-of-the-Art Approaches

In the previous section (5-1-1), after analyzing the results of the proposed algorithm on 10 valid UCI data sets, we found that the proposed method using opposition-based learning method of whale optimization algorithms and flower pollination from Categorization accuracy performs better in all data sets and has competitiveness in terms of the number of selected features compared to the whale optimization algorithms and flower pollination. In this sub-section, we compared the performance of the proposed method in addition to comparing the optimal whale-making algorithms and flower pollination with powerful extra-innovative methods such as genetic algorithm, particle density algorithm and bat algorithm. To show how the proposed method works compared to other powerful methods. In the experiment in this section, the parameters of all comparison algorithms are set with a maximum number of 80 repetitions and a population size of 10. Table (4) shows the results of comparing the proposed method with whale optimization algorithms, genetics, particle density, bat and flower spraying in terms of classification accuracy.

Table 4: Comparison between the proposed approaches and the state-of-the-art approaches in terms of classification accuracy.

Dataset	BAT	GA	PSO	WOA	FPA	HWOAFPA
Boold	0.6658	0.7406	0.7406	0.7406	0.6791	0.7406
BreastEW	0.9439	0.9719	0.9684	0.9614	0.9509	0.9649
CMC	0.3948	0.4858	0.4980	0.4980	0.4668	0.5061
Glass	0.9720	0.9813	0.9813	0.9813	0.9720	0.9813
Heart	0.6667	0.8074	0.7852	0.7852	0.5704	0.7852
Ionosphere	0.8466	0.9261	0.9261	0.9091	0.8523	0.9318
lymphography	0.5946	0.8378	0.8378	0.8108	0.6351	0.8378
Spect	0.6045	0.7761	0.6866	0.7687	0.5970	0.7463
Wine	0.8090	0.9438	0.9551	0.9663	0.6742	0.9488
Zoo	0.8627	0.9117	0.9412	0.9216	0.8235	0.9412

As shown in Table 4 results, the proposed algorithm is highly functional in terms of accuracy. In addition, the proposed algorithm performs better than all other methods in all data sets, except for three data sets, in which GA and other methods perform slightly better than the proposed method, and the proposed method ranks second. has it. Table 5 shows the results of comparing the proposed method with the optimized algorithms of whale, genetics, particle density, bat and flower spraying in terms of average selected features.

Table 5: Comparison between the proposed approaches and the state-of-the-art approaches in terms of average number of selected attributes.

Dataset	BAT	GA	PSO	WOA	FPA	HWOAFPA
Boold	2.30	1.00	1.71	1.00	1.90	1.00
BreastEW	11.7	12.5	16.6	18.0	20.0	12.1
CMC	4.00	3.00	3.61	4.00	6.00	4.00
Glass	3.30	2.00	5.69	3.00	6.10	2.00
Heart	5.00	5.00	6.27	5.00	8.30	4.30
Ionosphere	13.6	15.0	15.6	8.00	21.9	4.90
lymphography	5.40	8.90	8.46	3.00	12.1	11.00
Spect	8.10	10.0	12.0	4.00	14.6	4.00
Wine	4.00	5.00	6.90	7.00	8.20	5.00
Zoo	4.80	6.00	8.49	5.00	9.90	6.00

As shown in the results of Table (5), the proposed algorithm has a good performance in terms of the average of the selected features compared to other advanced super-innovative algorithms. In addition, the proposed algorithm performs better than all other methods in most data sets, except for a few data sets in which other methods perform better with a small difference from the proposed method, and the proposed method is in second place.

In this section, also for the convergence of the objective function defined in Equation (15), the convergence rate of each of the whale optimization algorithms, flower pollination, genetic algorithm, particle density algorithm and bat algorithm and finally the proposed method on 10 We have compared the valid UCI data set. To show how the proposed method works compared to other methods in terms of the degree of convergence of the objective function. In the experiment in this section, the parameters of all comparison algorithms are set with a maximum number of repetitions of 20 and a population size of 10. The results of this experiment are from Figures 9 to 14, respectively, and each of these figures is given to show the convergence of the algorithms on 10 data sets. Figure 9 shows the results of comparing the proposed method with the optimized algorithms of whale, genetics, particle density, bat and flower spraying in terms of the convergence of the objective function on the Boold data set.

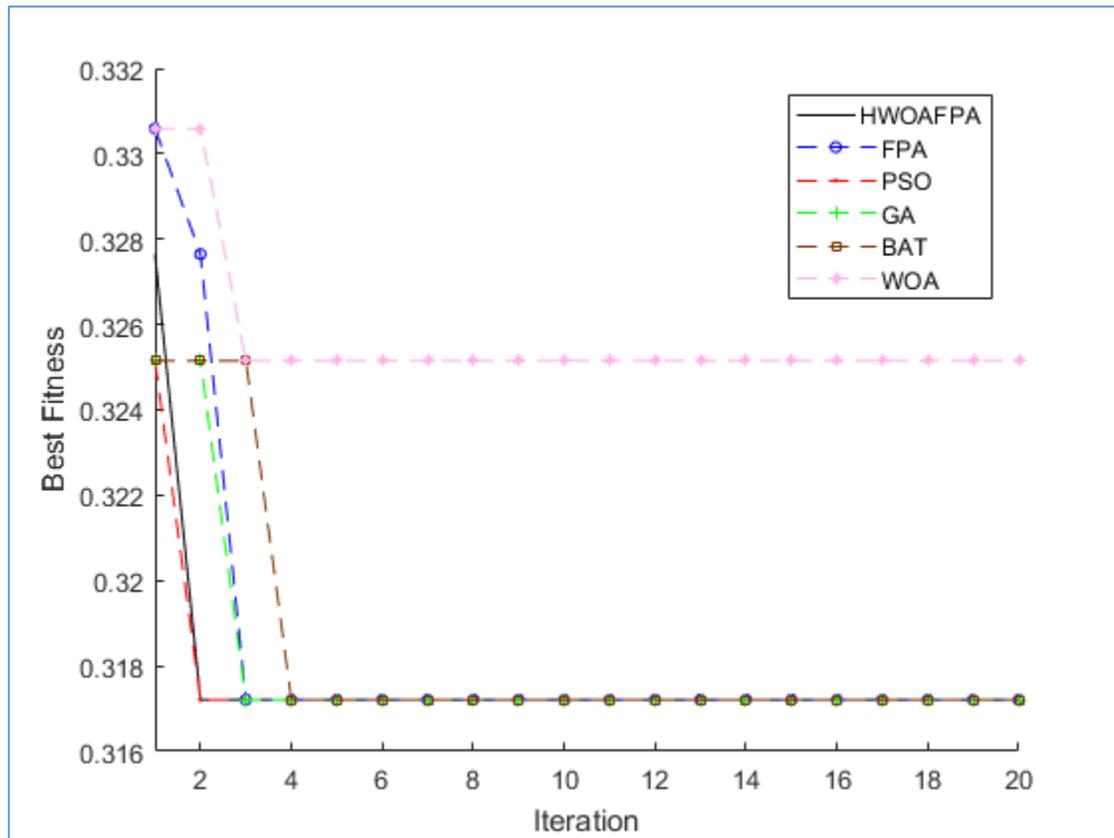


Fig. 9. Comparison between the proposed approaches and the state-of-the-art approaches in terms of convergence- Dataset Boold

As shown in Figure 9, the proposed algorithm and other comparative algorithms have almost the same results. The reason for the closeness of the results in this form is that the Boold data set has 4 features, and each of the algorithms, with multiple executions, ultimately achieves the optimal answer. Therefore, the results of the implementation on the Boold data set by the proposed method and the optimization algorithms of whale, genetics, particle density, bat and flower spraying show that due to the small size of the data set, all the algorithms compared in this article have similar results. Figure (10) shows the results of comparing the proposed method with the optimization algorithms of whale, genetics, particle density, bat and flower spraying in terms of the convergence of the objective function on the

BreastEW

data

set.

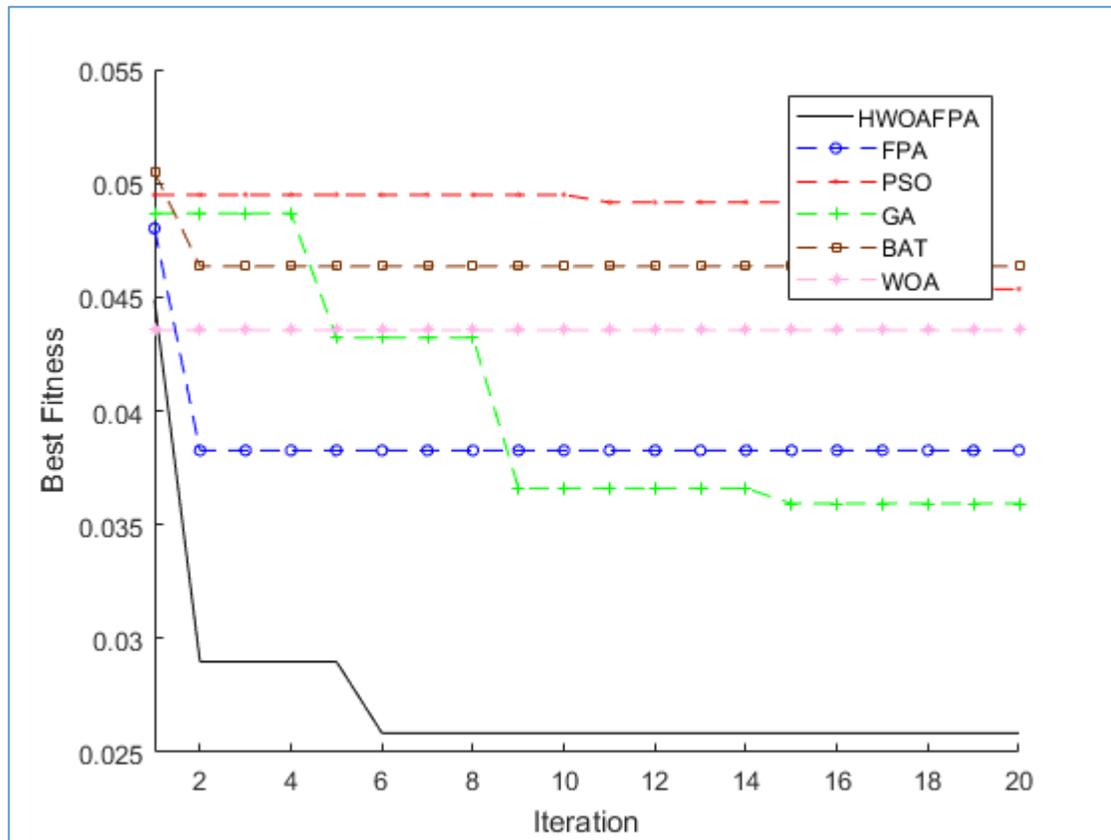


Fig. 10. Comparison between the proposed approaches and the state-of-the-art approaches in terms of convergence-Dataset BreastEW

As shown in Figure 10, the proposed algorithm and the ratio of other comparison algorithms have had a high performance in the convergence of the objective function, which is the high performance of the proposed algorithm because the BreastEW data set has more features. Each time the proposed algorithm has been able to use the opposite space of solutions, it is much stronger than other algorithms. Therefore, the results of the BreastEW data set by the proposed method and the optimized algorithms of whale, genetics, particle density, bat and flower spraying show that due to the large data set, comparative algorithms have reduced their efficiency but the proposed algorithm Due to the use of the opposite space of solutions in this article, the results are much better. Figure 11 shows the results of comparing the proposed method with the optimized algorithms of whale, genetics, particle density, bat and flower spraying in terms of the convergence of the objective function on the CMC data set.

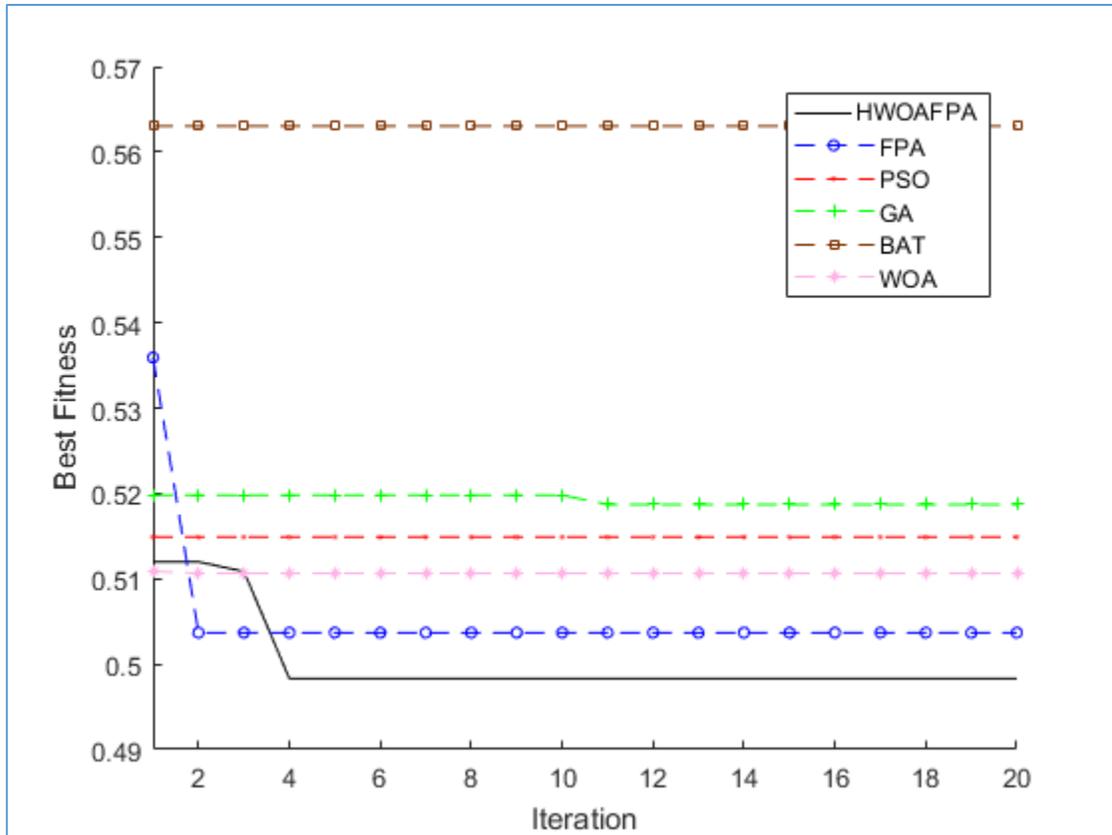


Fig. 11. Comparison between the proposed approaches and the state-of-the-art approaches in terms of convergence-Dataset CMC

As shown in Figure 11 (11), the proposed algorithm and the ratio of other comparative algorithms performed better in the convergence of the objective function. In this experiment, the performance of the proposed algorithm is such that the proposed algorithm has provided better results than the genetic algorithms, particle density, bat at the beginning, and has been able to repeat more using solutions and the opposite space of solutions in The final iterations of all comparative algorithms yield better results. Figure (12) shows the results of comparing the proposed method with the optimized algorithms of whale, genetics, particle density, bat and flower spraying in terms of the convergence of the objective function on the Glass data set.

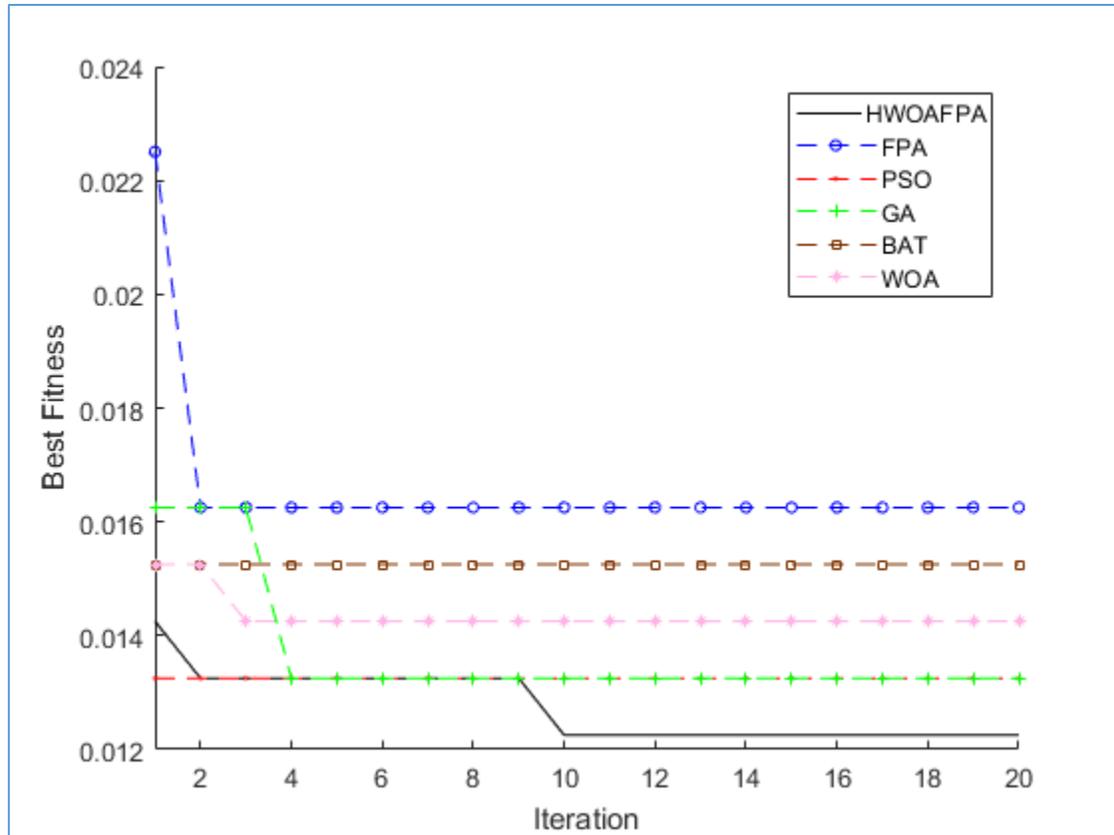


Fig. 12. Comparison between the proposed approaches and the state-of-the-art approaches in terms of convergence-Dataset Glass

As shown in the results of Figure (12), the proposed algorithm and the ratio of other comparative algorithms have performed relatively better in the convergence of the objective function. In this experiment, the performance of the proposed algorithm is such that the proposed algorithm has provided acceptable results at the beginning compared to the whale optimization and flower spraying algorithms, and has been able to replicate more using solutions and the opposite space of solutions. Get better results in the final iterations of all comparison algorithms. Figure 13 shows the results of comparing the proposed method with the optimized algorithms of whale, genetics, particle density, bat and flower spraying in terms of the convergence of the objective function on the Heart data set.

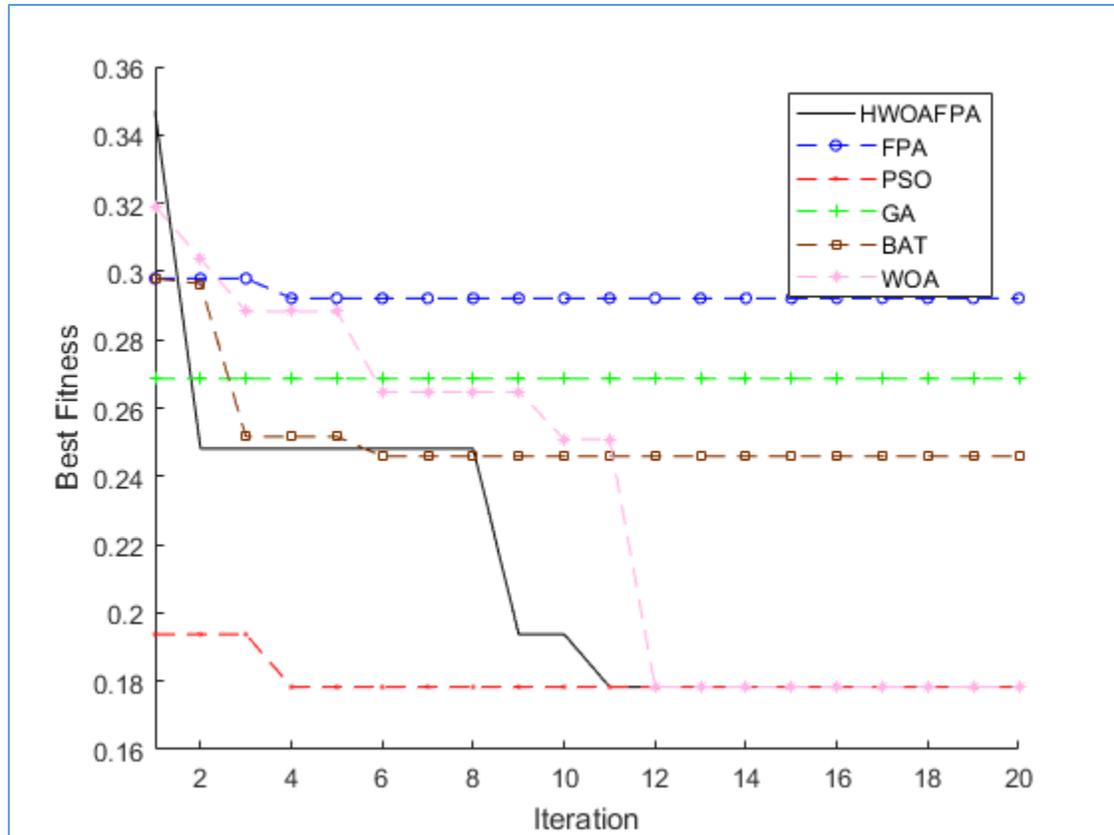


Fig. 13. Comparison between the proposed approaches and the state-of-the-art approaches in terms of convergence-Dataset Heart

As shown in Figure 13, the proposed algorithm and the ratio of other comparative algorithms performed similarly to whale optimization algorithms and particle density. Because the Heart data set is a median type of data set feature, it has made it possible for some algorithms such as the whale optimization algorithm and particle density to have a high performance in converging the objective function. Figure (14) shows the results of comparing the proposed method with the optimized algorithms of whale, genetics, particle density, bat and flower spraying in terms of the convergence of the objective function on the Ionosphere data set.

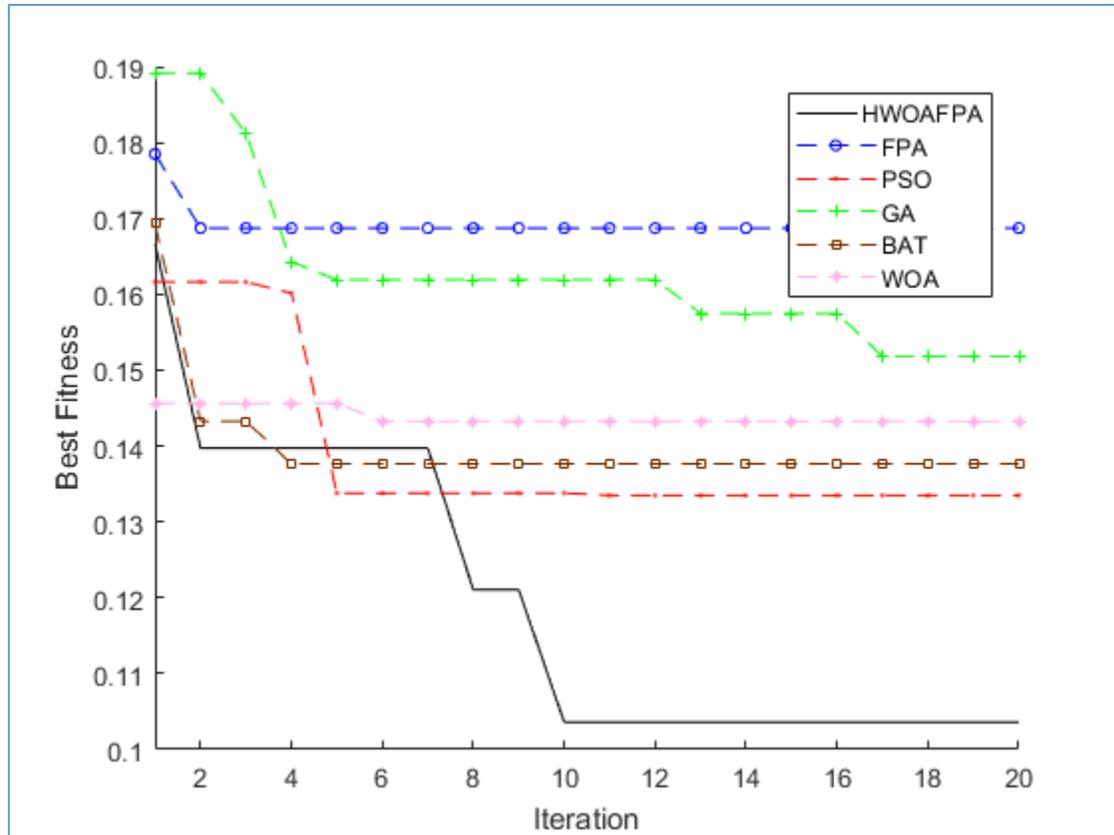


Fig. 14. Comparison between the proposed approaches and the state-of-the-art approaches in terms of convergence-Dataset Ionosphere

As shown in Figure 14, the proposed algorithm and the ratio of other comparative algorithms have had a high performance in the convergence of the objective function, which is the high performance of the proposed algorithm because the Ionosphere data set has more features. Each time the proposed algorithm has been able to use the opposite space, the solutions are much stronger than other algorithms. Therefore, the results of Ionosphere data collection by the proposed method and the optimized algorithms of whales, genetics, particle density, bats and flower spraying show that due to the large data set, comparative algorithms have reduced their efficiency but the proposed algorithm. Due to According to the results reported in these two sub-sections (5.1.1) and (5.1.2), a significant advantage of the proposed algorithm in selecting fewer features and high classification accuracy and convergence of the objective function compared to other meta-heuristic algorithms in comparison. It was found. Because we managed to provide a powerful algorithm with convergence speed and high accuracy to solve the problem of optimizing feature selection by creating a combination and using a learning method based on opposition.

5.2. Email Spam Detection

In the previous section (5-1), we implemented our proposed algorithm on 10 valid UCI data sets, the results showed that the proposed algorithm excelled in selecting fewer features and classification accuracy than other super-innovative algorithms. It was significant. Therefore, due to the strong results of the proposed algorithm on the valid UCI data set, it motivated us to implement our proposed algorithm on the spam email data set as well, in order to gain more confidence in the high performance of the proposed algorithm. To perform this experiment, we used the valid Spam email data set, which we will describe in more detail in sections (5-2-1). We distributed all the data using the classification sampling method to the test and training data set and tried to include spam and non-spam emails in both test and training sets to better evaluate the performance and efficiency of the proposed algorithm and other algorithms.

Let's have comparisons in this section. We also divided the spam data set into two parts, training and testing, for testing in this section.

- 70% of the entire dataset was used for training and construction of the proposed implementation model
- 30% of the remaining dataset was used for testing and validating the model.

5.2.1. Spam base dataset analysis

The Corpus bench brand is emailed from the spam database. At the time of receiving this spam message, it contained 4,601 messages, of which 1,813 (39%) were known as spam messages and 2,788 (61%) were known and purchased as non-spam messages. Were [44-49]. The command helped the non-spam message. This is taken from a single mailbox. Unlike most failures in their raw form, the purchase of this carcass has already been pre-processed. Accidents or characteristics are presented as 58-dimensional vectors. In a set of 57 attributes, 48 of the 48 attributes are displayed with words that consist of the main messages, which have no pause or stop list, and are placed in the list of the most unbalanced words for the class spam. The remaining 6 attributes are the percentage of specific characters ";", "" (" "" "" "" "" "!", "\$" And "#", which allows spam or non-spam to be transmitted through agencies 1 and 0 All data, including the database, is a good test bed.

5.2.2. Experimental results and discussion

In this section, we will show the results of the implementation of the proposed algorithm and other comparative algorithms on the data set of spam emails, and then we will discuss the results obtained. We divided the spam email data collection into test data set (30%) and training (70%), and in both test and training data sets, spam and non-spam emails are placed to better evaluate the performance and efficiency of the algorithm. We have suggestions and other comparative algorithms in this section. We also considered the initial population of all algorithms to be 10 and the number of iterations to be 100-1. In this section, we will first test the proposed algorithm with the WOA, FPA algorithm in terms of classification accuracy and average classification accuracy with different repetitions, and then in the next step we will propose the proposed algorithm and other comparative algorithms on the data set of spam emails from in terms of classification accuracy, we experiment with different iterations.

In the first experiment of this sub-section, the performance of WOA, FPA and HWOAFPA in terms of accuracy of detecting spam emails is shown. The parameters are set to a maximum number of repetitions of 1-100 and a population size of 10.

The results of the first experiment in this subsection, show that the proposed algorithm has achieved much better results and accuracy than whale optimization and flower spraying algorithms in email detection by increasing the number of iterations. The proposed algorithm works in the first repetitions like other powerful algorithms, but as the number of repetitions in increases, it can be seen that it has been able to excel due to its greater use of opposition-based learning methods to create solutions in the opposite search space. Show the ratio of whale optimization algorithms and flower pollination.

For further testing in this section, in addition to the experimental accuracy classification, we examined the proposed algorithm in terms of the average accuracy of spam email detection, which includes the entire population of the proposed algorithms and WOA and FPA. Therefore, the WOA, FPA and HWOAFPA algorithms are shown in terms of the average accuracy of spam email detection. The parameters are set to a maximum number of repetitions of 1-100 and a population size of 10.

The results of the second experiment in this subsection, show that by increasing the number of repetitions in the proposed algorithm, much better results are obtained from whale optimization and flower pollination algorithms in terms of average accuracy of spam email detection accuracy. Is. This experiment proves that the proposed algorithm in the number of repetitions changes its entire population and improves all the solutions available in the search space. Now, after comparing the proposed algorithm with whale optimization and flower pollination algorithms and showing the superiority of the proposed algorithm in terms of classification accuracy and average classification accuracy, we will try to further evaluate the proposed algorithm with other proposed algorithms. We test meta-initiatives such as genetic algorithms and particle and bat crowding in terms of classification accuracy with different repetitions.

The results of the third experiment in this subsection, show that the proposed algorithm has stronger results than comparative algorithms such as genetic algorithms, particle density, bat, whale optimization and flower pollination in terms of accuracy in detecting spam emails. Has achieved. The results of these three experiments in the discussion of spam email detection proved that the combination was well done in the proposed method. Using the opposition-based learning method in the combined mode, we were able to provide a powerful algorithm for detecting spam emails.

6. Conclusion

A feature selection method should search for the best subset of features among all possible subsets in the data set and extract these properties based on a specific evaluation criterion. Applying feature selection in the data set before the learning process plays an essential role in improving the efficiency of the categorization process. Finding this optimal subset can be considered as an optimization problem and using random, innovative and meta-innovative search methods to find these optimal or semi-optimal subsets. The important point is that when the search space is large, choosing a subset of features using traditional and innovative optimization methods is not effective and efficient, and meta-innovative algorithms as the most suitable option to solve this problem by many researchers so far. Provided. Innovative algorithms may also be combined to increase the convergence rate and accuracy of the production solution. In this paper, we have tried to solve the problem of optimization of feature selection using the natural processes of whale optimization and flower pollination algorithms, and on the other hand, we have used the opposition-based learning method to ensure the convergence speed and accuracy of the proposed algorithm. In fact, in the proposed method, whale optimization algorithms use the bait siege process, bubble attack method and bait search to create solutions in their search space and try to improve the solutions to the feature selection problem, along with this algorithm. The pollen pollination algorithm with two global and local search processes improves the problem-solving feature in contrasting solutions with the whale optimization algorithm. In fact, we used both search space solutions and contrasting search space solutions, all possible solutions to the feature selection problem. To evaluate the performance of the proposed algorithm, we performed the experiments in two stages. In the first phase, experiments were performed on 10 sets of data selection features from the UCI data repository. In the second step, we tried to test the performance of the proposed algorithm by detecting spam emails. The results obtained from the first step show that the proposed algorithm, by running on 10 UCI data sets, has been able to be more successful in terms of average selection size and classification accuracy than other basic meta-heuristic algorithms. Also, the results obtained from the second step show that the proposed algorithm has been able to perform spam emails more accurately than other modern meta-heuristic algorithms in terms of accuracy.

7. References

1. Mafarja, M., et al., *Evolutionary Population Dynamics and Grasshopper Optimization Approaches for Feature Selection Problems*. Knowledge-Based Systems, 2017.
2. Mafarja, M.M. and S. Mirjalili, *Hybrid Whale Optimization Algorithm with simulated annealing for feature selection*. Neurocomputing, 2017. **260**: p. 302-312.
3. Han, J., J. Pei, and M. Kamber, *Data mining: concepts and techniques*. 2011: Elsevier.
4. Aghdam, M.H., N. Ghasem-Aghaee, and M.E. Basiri, *Text feature selection using ant colony optimization*. Expert systems with applications, 2009. **36**(3): p. 6843-6853.
5. Zorarpacı, E. and S.A. Özel, *A hybrid approach of differential evolution and artificial bee colony for feature selection*. Expert Systems with Applications, 2016. **62**: p. 91-103.
6. Huang, J., Y. Cai, and X. Xu, *A hybrid genetic algorithm for feature selection wrapper based on mutual information*. Pattern Recognition Letters, 2007. **28**(13): p. 1825-1844.
7. Kabir, M.M., M. Shahjahan, and K. Murase, *A new hybrid ant colony optimization algorithm for feature selection*. Expert Systems with Applications, 2012. **39**(3): p. 3747-3763.
8. Khalid, S., T. Khalil, and S. Nasreen. *A survey of feature selection and feature extraction techniques in machine learning*. in *Science and Information Conference (SAI), 2014*. 2014. IEEE.

9. Xue, B., et al., *A survey on evolutionary computation approaches to feature selection*. IEEE Transactions on Evolutionary Computation, 2016. **20**(4): p. 606-626.
10. Rajamohana, S. and K. Umamaheswari, *Hybrid approach of improved binary particle swarm optimization and shuffled frog leaping for feature selection*. Computers & Electrical Engineering, 2018.
11. Xue, B., M. Zhang, and W.N. Browne, *Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms*. Applied Soft Computing, 2014. **18**: p. 261-276.
12. Wang, Z., M. Li, and J. Li, *A multi-objective evolutionary algorithm for feature selection based on mutual information with a new redundancy measure*. Information Sciences, 2015. **307**: p. 73-88.
13. Emary, E., H.M. Zawbaa, and A.E. Hassanien, *Binary ant lion approaches for feature selection*. Neurocomputing, 2016. **213**: p. 54-65.
14. Nakamura, R.Y.M., et al., *Binary bat algorithm for feature selection*, in *Swarm Intelligence and Bio-Inspired Computation*. 2013, Elsevier. p. 225-237.
15. Zhang, L., et al., *Feature selection using firefly optimization for classification and regression models*. Decision Support Systems, 2017.
16. Mafarja, M. and S. Mirjalili, *Whale optimization approaches for wrapper feature selection*. Applied Soft Computing, 2018. **62**: p. 441-453.
17. Kabir, M.M., M. Shahjahan, and K. Murase, *A new local search based hybrid genetic algorithm for feature selection*. Neurocomputing, 2011. **74**(17): p. 2914-2928.
18. Sayed, S.A.-F., E. Nabil, and A. Badr, *A binary clonal flower pollination algorithm for feature selection*. Pattern Recognition Letters, 2016. **77**: p. 21-27.
19. Liao, T. and R. Kuo, *Five Discrete Symbiotic Organisms Search Algorithms for Simultaneous Optimization of Feature Subset and Neighborhood Size of KNN Classification Models*. Applied Soft Computing, 2017.
20. Tawhid, M.A. and K.B. Dsouza, *Hybrid Binary Bat Enhanced Particle Swarm Optimization Algorithm for solving feature selection problems*. Applied Computing and Informatics, 2018.
21. Shunmugapriya, P. and S. Kanmani, *A hybrid algorithm using ant and bee colony optimization for feature selection and classification (AC-ABC Hybrid)*. Swarm and Evolutionary Computation, 2017. **36**: p. 27-36.
22. Oh, I.-S., J.-S. Lee, and B.-R. Moon, *Hybrid genetic algorithms for feature selection*. IEEE Transactions on pattern analysis and machine intelligence, 2004. **26**(11): p. 1424-1437.
23. Khushaba, R.N., et al. *A combined ant colony and differential evolution feature selection algorithm*. in *International Conference on Ant Colony Optimization and Swarm Intelligence*. 2008. Springer.
24. Olabiyisi, S., et al., *Hybrid metaheuristic feature extraction technique for solving timetabling problem*. Int. J. Sci. Eng. Res, 2012. **3**(8): p. 1-6.
25. Canuto, A.M. and D.S. Nascimento. *A genetic-based approach to features selection for ensembles using a hybrid and adaptive fitness function*. in *Neural Networks (IJCNN), The 2012 International Joint Conference on*. 2012. IEEE.
26. Akila, M.M., *Hybrid local feature selection in DNA analysis based cancer classification*. Indian Journal of Computer Science and Engineering, 2012. **3**(3): p. 470-475.
27. Babatunde, R., S. Olabiyisi, and E. Omidiora, *Feature dimensionality reduction using a dual level metaheuristic algorithm*. optimization, 2014. **7**(1).
28. Hasani, S.R., Z.A. Othman, and S.M.M. Kahaki, *Hybrid feature selection algorithm for intrusion detection system*. Journal of Computer Science, 2014. **10**(6): p. 1015.
29. Kashef, S. and H. Nezamabadi-pour, *An advanced ACO algorithm for feature subset selection*. Neurocomputing, 2015. **147**: p. 271-279.
30. Nekkaa, M. and D. Boughaci, *Hybrid harmony search combined with stochastic local search for feature selection*. Neural Processing Letters, 2016. **44**(1): p. 199-220.
31. Mirjalili, S. and A. Lewis, *The whale optimization algorithm*. Advances in Engineering Software, 2016. **95**: p. 51-67.
32. Glover, B., *Understanding flowers and flowering second edition*. 2014: OUP Oxford.

33. Waser, N.M., *Flower constancy: definition, cause, and measurement*. The American Naturalist, 1986. **127**(5): p. 593-603.
34. Pavlyukevich, I., *Lévy flights, non-local search and simulated annealing*. Journal of Computational Physics, 2007. **226**(2): p. 1830-1844.
35. Yang, X.-S. *Flower pollination algorithm for global optimization*. in *International conference on unconventional computing and natural computation*. 2012. Springer.
36. Rodrigues, D., et al., *Binary flower pollination algorithm and its application to feature selection*, in *Recent advances in swarm intelligence and evolutionary computation*. 2015, Springer. p. 85-100.
37. Rahnamayan, S., H.R. Tizhoosh, and M.M. Salama, *Opposition-based differential evolution*. IEEE Transactions on Evolutionary computation, 2008. **12**(1): p. 64-79.
38. Shaw, B., V. Mukherjee, and S. Ghoshal, *A novel opposition-based gravitational search algorithm for combined economic and emission dispatch problems of power systems*. International Journal of Electrical Power & Energy Systems, 2012. **35**(1): p. 21-33.
39. Rahnamayan, S., H.R. Tizhoosh, and M.M. Salama, *Opposition versus randomness in soft computing techniques*. Applied Soft Computing, 2008. **8**(2): p. 906-918.
40. Altman, N.S., *An introduction to kernel and nearest-neighbor nonparametric regression*. The American Statistician, 1992. **46**(3): p. 175-185.
41. Mabodi, K., Yusefi, M., Zandiyan, S., Irankhah, L., & Fotohi, R. (2020). Multi-level trust-based intelligence schema for securing of internet of things (IoT) against security threats using cryptographic authentication. The Journal of Supercomputing, 1-25.
42. Seyedi, B., & Fotohi, R. (2020). NIASHPT: a novel intelligent agent-based strategy using hello packet table (HPT) function for trust Internet of Things. The Journal of Supercomputing, 1-24.
43. Fotohi, R. (2020). Securing of Unmanned Aerial Systems (UAS) against security threats using human immune system. Reliability Engineering & System Safety, 193, 106675.
44. Fotohi, R., Firoozi Bari, S., & Yusefi, M. (2020). Securing Wireless Sensor Networks Against Denial-of-Sleep Attacks Using RSA Cryptography Algorithm and Interlock Protocol. International Journal of Communication Systems, 33(4), e4234.
45. Asuncion, A. and D. Newman, *UCI machine learning repository*. 2007.
46. Merz, C.J. and P.M. Murphy, *{UCI} Repository of machine learning databases*. 1998.
47. Friedman, J., T. Hastie, and R. Tibshirani, *The elements of statistical learning*. Vol. 1. 2001: Springer series in statistics New York.
48. Mark Hopkins, E.R., George Forman, Jaap Suermondt *UCI Machine Learning Repository: Spambase Data Set*, in *Hewlett-Packard Labs*. 1999.
49. Koprinska, I., et al., *Learning to classify e-mail*. Information Sciences, 2007. **177**(10): p. 2167-2187.