JACET

J. ADV COMP ENG TECHNOL, 7(1) Winter 2021

An Optimization-based Learning Black Widow Optimization Algorithm for Text Psychology

***Ali Hosseinalipour*[1], *Farhad Soleimanian Gharehchopogh*[2], *Mohammad Masdari*[3], *Ali Khademi*[4]**
1- Department of Computer Engineering, Urmia Branch, Islamic Azad University, Urmia, IRAN.
2- Department of Computer Engineering, Urmia Branch, Islamic Azad University, Urmia, IRAN. (bonab.farhad@gmail.com)
3- Department of Computer Engineering, Urmia Branch, Islamic Azad University, Urmia, IRAN.
4- Department of Psychology Science, Urmia Branch, Islamic Azad University, Urmia, IRAN.

Received (2021-01-03) Accepted (2021-09-18)

Abstract: *In recent years, social networks' growth has increased these networks' content. Therefore, text mining methods became important. As part of text mining, Sentiment analysis means finding the author's perspective on a particular topic. Social networks allow users to express their opinions and use others' opinions in other people's opinions to make decisions. Since the comments are in the form of text and reading them is time-consuming. Therefore, it is essential to provide methods that can provide us with this knowledge usefully.*

Black Widow Optimization (BWO) is inspired by black widow spiders' unique mating behavior. This method involves an exclusive stage, namely, cannibalism. For this reason, at this stage, species with an inappropriate evaluation function are removed from the circle, thus leading to premature convergence.

In this paper, we first introduced the BWO algorithm into a binary algorithm to solving discrete problems. Then, to reach the optimal answer quickly, we base its inputs on the opposition. Finally, to use the algorithm in the property selection problem, which is a multi-objective problem, we convert the algorithm into a multi-objective algorithm. The 23 well-known functions were evaluated to evaluate the performance of the proposed method, and good results were obtained. Also, in evaluating the practical example, the proposed method was applied to several emotion datasets, and the results indicate that the proposed method works very well in the psychology of texts.

Keywords: *text psychology, meta-heuristic algorithm, feature selection, black widow optimization algorithm.*

How to cite this article:
Ali Hosseinalipour, Farhad Soleimanian Gharehchopogh, Mohammad Masdari, Ali Khademi. An Optimization-based Learning Black Widow Optimization Algorithm for Text Psychology. J. ADV COMP ENG TECHNOL, 7(1) Winter 2021 : 81-92

I. INTRODUCTION

Nowadays, a large amount of textual data is available for processing, and this data is increasing every year. An essential part of the way human beings gather information is to understand what other human beings think. An essential part of the information that people use during their decision-making process is always based on the answer to the question, "What do others think?" With the accessibility and popularity of rich sources of ideas such as online review sites, personal blogs, and social networks, new opportunities and challenges have been created in this area. People can now use information technology to analyze the feelings of others. Hence, given the growing interest in systems that can understand the thoughts and opinions of others. It should be noted that the only motivation for people to explore ideas online is marketing business products and services. Access to political information, for example, is another critical application in this area.

The optimization process is to find the

best arrangement of organisms that seek limited resources with pre-defined constraints. This process can be used in several research fields such as health, engineering, mathematics, economics, linguistics, and science to optimize (minimize or maximize) their purpose. There are two types of optimization methods based on determinism and approximation[1-3]. Traditionally, certainty-based methods have been used to deal with smaller size optimization problems and less complexity. Although they can find an exact solution to the optimization problem, they suffer from various problems, including that they cannot solve NP problems. They can quickly get stuck in a local optimum[4-6]. Therefore, they are inefficient in dealing with real-world problems. As a result, optimization research communities turn their attention to using approximation methods to solve their optimization problems.

Due to the increase in the volume and dimensions of information, feature selection is essential when using machine learning and data mining methods. Feature selection is a practical and fundamental step considered a prerequisite for classification methods[7-9]. Large data sets for teaching a category may lead to a problem called over-fitting in learning methods. The over-fitting problem reduces the model's generalizability and reduces the accuracy of classification methods for new test samples. Moreover, a large data set requires more processing time to build the model from the training and testing data set [10]. As a result, feature selection aims to simplify and improve the data set's quality by selecting essential and critical features[11]. Also, feature selection can better understand the domain and, according to some criteria, retain only better and more appropriate features to describe the inherent patterns in the data and help reduce the effects of dimensions[12].

This paper's motivation is to propose a suitable algorithm in the psychology of texts, for which we have improved and used the BWO algorithm. To do this, we first use the BWO algorithm to improve its speed based on the opposition. Then, we convert the improved algorithm to a binary algorithm because we want to use this method in the discrete Feature selection problem. We turn it into a multi-objective feature because it is also a matter of selecting a multi-objective algorithm. Finally, we use this algorithm to classify psychological texts. In order to evaluate the performance of opposition-based, discrete, and multi-objective BWO, 23 well-known functions are used for evaluation criteria. The results show that it escapes the local optimum and balances the exploitation and exploration stages compared to the other algorithms studied. For added validity, the proposed solution is used to analyze the author's feelings of the text as a practical example. The simulation results show that the proposed solution is efficient.

The advantages of the proposed algorithm are: I- They converge quickly. II- Because the inputs are based on contradiction, they reach the optimal answer faster. III- Also, compared with other methods in different evaluations, the proposed algorithm works better.

The remaining sections of this paper are as follows: Related work is provided in Section II. Materials and methods are introduced in section III. The proposed Multi-objective Opposition-based Binary BWO (MOBBWO) algorithm is introduced in Section IV. The performance of the proposed algorithm and evaluation and analysis are presented in Section V. Section VI provides a practical example of the proposed method. Finally, the conclusion is shown in the last section.

II. AN OVERVIEW OF PREVIOUS METHODS

Extensive research has recently been conducted on the topic of Sentiment analysis and belief. Around 2001, there was widespread awareness of Sentiment analysis and ideology's research issues, and subsequently, thousands of papers in the field were published. In 2001, Das, Chen, and Tong (in separate papers) analyzed market sentiment. Subsequently, in 2002, papers published by Turney and Pang et al. At the annual meeting of the Association for Computational Linguistics (ACL) and the annual conference of Empirical Methods in Natural Language Processing (EMNLP) these words were used[13]. Nasukawa and Yi published a paper in 2003 entitled Sentiment analysis: Capturing favorability using natural language processing[14]. Some meta-heuristic algorithms have been used in feature selection in text psychology, with TABLE I showing some of the other meta-heuristic algorithms used in feature selection.

TABLE I
SEVERAL METHODS RELATED TO FEATURE SELECTION

Ref.	method	domain	description
[15]	A review of feature extraction in sentiment analysis	Sentiment analysis	A brief review of feature selection techniques
[16]	A review of feature selection techniques in bioinformatics	Bioinformatics	The primary classification of feature selection techniques
[17]	A Comprehensive Analysis of Nature-Inspired Meta-Heuristic Techniques for Feature Selection Problem	Feature selection problem	A researcher who tends to design or analyze the performance of meta-innovative divergence methods in solving feature selection problem
[18]	An efficient binary social spider algorithm for the feature selection problem	Binary optimization	Social spider algorithm for the feature selection problem

Before addressing the proposed method, we must recall the challenges in implementing text author emotion analysis, none fully addressed in natural language processing. We should try to overshadow these cases in the proposed method. Some of these challenges are:

1. The user does not express his feelings directly and may express his opinion in various ways (negative actions or attributes). Consecutive sentences may complement each other.

"I was not feeling well at all today, or if you think I was feeling well, you are wrong."

2. The user may use sarcasm. Each of these items has an entirely different meaning in the sentence structure from the fundamental concept.

"I was so happy today that a lamb in the slaughterhouse is happy!"

"This computer is not as usable as a piece of brick."

3. Users often use complex sentence structures that do not necessarily fall within the general framework of language structure due to their informality.

4. For the above reasons, a word-for-word review can lead to erroneous results.

In the proposed method, we will use a different combination of methods. This method will be done in three steps, the first of which is preprocessing and normalization. The second step is to select the compelling features with the least number in the category of emotions, which will be done using the ultra-innovative algorithm of BWO. The last step is the labeling and categorization of emotions. This paper used BWO, an exemplary method compared to other meta-innovative methods because it has a high convergence speed. Moreover, it performs well in feature selection and can therefore overshadow existing challenges and eliminate a percentage of them.

III. MATERIAL AND METHODS

In this section, the BWO algorithm will be described, which is used as the basic algorithm, then the data set used in this paper is introduced.

1. The BWO algorithm

The BWO algorithm, like other evolutionary algorithms, begins with the initial population of spiders so that each spider represents a potential solution. These early spiders, in pairs, try to reproduce the new generation. The black widow eats the male spider during or after mating. He then carries the stored sperm into his sperm cavities and releases them into the egg sacs. Eleven days after fertilization, the spiders emerge from the egg sacs. They live on the mother net for several days to a week, during which time it is observed that weaker siblings are eaten. The main steps of the algorithm are described below[19].

1-1- The initial population

In the BWO, it is called a "widow." The potential solution to any problem is considered as a black widow spider. Each black widow spider represents the values of the problem variables. In this method, the structure must be considered as an array to solve the benchmark functions. In the next N_{Var} optimization problem, a widow represents an $N_{Var} \times 1 array$ that represents the solution to the problem. This array is defined as an Equation (1):

$$widow = [x_1, x_1, \dots, x_{(N_{Var})}], \qquad (1)$$

In Equation (1), widow means widow spiders and a subsequent N_{Var} array and x is the position of each widow spider. The widow evaluation function is obtained by evaluating the performance of the f evaluation function in widow a from$(x_1, x_1, \dots, x_{N_{Var}})$. Therefore, the evaluation function is calculated as Equation (2):

$$Fitness = f(widow) = (x_1, x_1, \dots, x_{(N_{Var})}) \quad (2)$$

In Equation (2), Fitness or f (widow) means the evaluation function of widow spiders. To start the optimization algorithm, a candidate widow matrix of size $N_{Var} \times N_{pop}$ with the initial population of spiders is generated. The parent pair is then randomly selected to perform the calving stage by mating, in which the black widow eats the male during or after.

1-2- Reproduction

Because these pairs are not independent, each pair is separated from the others in parallel and nature to reproduce the new generation. In the real world, every mate produces about 1,000 eggs, but in the end, some of the more robust baby spiders survive. Now, here in this algorithm for reproduction, an array called alpha must be created. First, the widow array is filled with random numbers, and then the children are generated using α with and Equation (3).

$$\begin{aligned} y_1 &= \alpha \times x_1 + (1-\alpha) \times x_2 \\ y_2 &= \alpha \times x_2 + (1-\alpha) \times x_1 \end{aligned} \quad (3)$$

In Equation (3), x_1 and x_2 are the parents, y_1 and y_2 are the children. This process is repeated for $N_{Var}/2$ times, while the randomly selected numbers should not be repeated. Eventually, children and mothers are added to an array and sorted by their evaluation function's value. According to the cannibalism ranking, some of the best people are added to the newly formed population. These steps apply to all pairs.

1-3- Cannibalism

Here we have three types of cannibalism. The first is sexual cannibalism, in which a black widow eats her husband during or after mating. In this algorithm, men and women are identified according to their organ evaluation function. Another type of cannibalism is sibling, in which healthy spiders eat their weaker siblings. In this algorithm, a cannibalism rating (CR) is set based on the number of survivors. In some cases, the third type of cannibalism is often seen, in which the baby eats its mother spider. The value of the limb evaluation function is used to determine solid or weak spiders[19].

1-4- Mutation

At this point, the Mutepop number is randomly selected from the population. Each of the selected solutions randomly exchanges two elements in the array. Muttpop is calculated with the mutation rate[19].

1-5- Convergence

Like other evolutionary algorithms, three-stop conditions can be considered[19]: (A) Pre-defined number of iterations. (B) Observe no change in the value of the best widow evaluation function for multiple iterations. (C) Achieve a certain level of accuracy.

1-6- Parameter setting

In the BWO algorithm, some parameters are necessary to achieve better results. These parameters include reproduction rate (PP), cannibalism rate (CR), and mutation rate (PM). The parameters must be adjusted appropriately to improve the algorithm's success in finding superior solutions. The better the adjusted parameters, the more likely you will jump out of any global optimized search space. Hence, the right amount of parameters can ensure balance control between the exploitation and exploration stages.

2. Data set

Two sets of data are used in this paper: the details of each are as follows.

2-1- ISEAR data set

This dataset was developed by a large group of psychologists around the world during the 1990s. This 40-feature dataset includes psychological questions designed under the supervision of Klaus and Harald Wallenboot. In this test, respondent students (both psychologists and non-psychologists) were asked to report situations where they experienced seven primary emotions (happiness, fear, anger, sadness, hatred, shame, and guilt). In each case, the questions included assessing the situation and how the respondents reacted. The final data set is based on reports of seven primary emotions involving approximately 3,000 respondents in 37 countries on five continents.

2-2- Sentiment polarity datasets v2.0

The name of this collection is Movie Review Data, the second version called Sentiment polarity datasets v2.0. It is a film analysis data set from an emotional perspective. This collection includes movie review texts with two positive and negative classes. The data set consists of 2000 texts with specific tags, including a thousand texts with a positive view and a thousand texts with an opposing view[20].

IV. THE PROPOSED METHOD

The psychology of texts is based on dynamic analysis and recognition of the author's behavior and is one of the new research areas that have recently been considered in some languages, especially English. One of the most critical challenges in distinguishing the author's mood and feeling from the text from other processing operations in natural language processing is that individual situations and their reflection in his speech and writing depend very strongly on the author's culture and nationality. However, the advantage that can be found in this category is that it is easy to identify, determine and compile the keywords, key sentences, and punctuation marks that determine these behavioral states in the relevant language, and based on that, the author's state at the time of writing the text recognized. In this paper, we first turn the BWO algorithm into a discrete algorithm. To solve the discrete problem, we can use feature selection, which is a discrete problem. Then turn it into a conflict-based algorithm to speed up the optimal answer. The resulting algorithm is then multi-objective because the feature selection problem is a multi-objective problem that pursues feature number and classification accuracy. Finally, it is used for the psychology of the text.

The structure of the proposed method is as follows: first, in subsection1, the BWO algorithm is binary, then in subsection 2, it is based on opposition to improve its speed, and in subsection 3, the algorithm obtained from the above steps is multi-algorithm The goal becomes. In Section VI, the resulting algorithm is used in text psychology.

1. Optimization-based learning for optimizing binary BWO based on sigmoid function

This section will introduce a new binary method for the BWO algorithm based on the sigmoid function. As stated in subsection 1 III, the BWO algorithm moves in continuous space, and therefore all the solutions available in the population of this algorithm include serial numbers. Given that it is a matter of selecting or not selecting a Feature, the new binary solution must contain the numbers 0 and 1, where 1 indicates selecting a feature for the new dataset and zero indicates the non-selection of a feature. To do this, we will use the sigmoid function or the same[21, 22] S-shaped to move the processes of the BWO algorithm in binary space. Therefore, in this proposed model, the sigmoid function is used to continuously change the position of the solutions in the BWO algorithm to binary mode as an Equation (4):

$$sg\left(BWO_i^d(t)\right) = \frac{1}{1 + e^{-BWO_i^d(t)}} \tag{4}$$

In Equation (4), BWO_i^d is the continuous value of solution i in the BWO algorithm's population in the d[th] dimension in iteration t. The output of the sigmoid transfer function is still in a continuous state between 0 and 1, so a threshold must be set to convert it to a binary value, which is the random threshold given in Equation (5) to convert the binary value solution to select the Feature in The sigmoid function is applied:

$$BWO_i^d(t+1) = \begin{cases} 0 & if\ rand < sg\left(BWO_i^d(t)\right) \\ 1 & if\ rand \geq sg\left(BWO_i^d(t)\right) \end{cases} \tag{5}$$

In Equation (5), BWO_i^d represents the i

solution's position in the BWO algorithm population in iteration t in the d dimension. Rand also represents a number between zero and one of the uniform distribution type. Thus, the solutions available in the BWO algorithm population are forced to move in a binary search space using Equations (4) and (5). Next, we place these relationships in more detail in the BWO algorithm.

2. opposition-based binary improved BWO algorithm

In the next step, the algorithm's inputs are based on contradiction, according to Equation 6. In methods based on opposition to each member of the original population, a contradictory member is also produced. If the conflicting member cost function is less than the original member cost function, it can be substituted; otherwise, we will continue. Therefore, its members and opposing members are evaluated simultaneously to proceed with those that are more appropriate. Assuming that x is the position of the spider between a and b, the opposition-based $\overline{X}$ is defined as Equation 6.

$$\overline{X} = a + b - X \qquad (6)$$

In Equation (6), X is the position of the spider between a and b, $\overline{X}$ is based on the opposition.

Given that feature selection is discrete, the BWO algorithm is inherently a continuous algorithm. Therefore, the mapping strategy is used to convert actual variables into correct variables. This operation is performed using the correct component function shown in Equation 6. In this formula, x is a fundamental variable between b and a, where b and a are two consecutive integers. This strategy can solve the problem of the continuity of the algorithm to become a discrete problem. Moreover, it is known as OBBWO after being the basis for opposing this algorithm.

3. Optimized binary BWO algorithm based on multi-objective conflict

This section describes the feature selection objective function for the proposed algorithm and other meta-heuristic algorithms in this paper. Feature selection can be considered a multi-objective optimization problem in which two conflicting goals are achieved, including the minimum number of features selected and higher categorization accuracy. Therefore, to define the feature selection problem's objective function, we need a classification algorithm and the simplest classification method, i.e., the KNN classifier. We also used this classification to define the objective function of the feature selection problem. Therefore, in the proposed method, we used the KNN classifier to evaluate the proposed algorithm's features and other algorithms more accurately. Each solution is evaluated based on the proposed multi-objective function, which depends on the KNN classifier. In the proposed multi-objective function to balance the number of features selected in each solution (minimum) and classification accuracy (maximum), the proportionality function in Equation (7) is used to evaluate a solution in any meta-heuristic algorithms.

$$Fitness = \alpha\gamma_R(D) + \beta\frac{|R|}{|N|} \qquad (7)$$

In Equation (6), $\alpha\gamma_R(D)$ indicates the classification error rate of a classification method. Also, |R| Multi-linear subset is selected, and |N| is the total number of features in the dataset. Also, the parameter α is the importance of category quality, and the parameter β is the length of the subset. The values of these two parameters are α ∈ [0, 1] and β = (1 -α). In this study, the initial value of α is 0.99; B's value will be 0.01. Considering that the main algorithm is called BWO, we have named it discrete BBWO, and based on the contradiction, we have named it OBBWO, and finally, for several purposes, we have called this method MOBBWO.

V. SIMULATION AND EVALUATION OF RESULTS

1. Results with 23 evaluation functions

Twenty-three standard benchmarks have been considered to evaluate the performance of the proposed method. All of these test functions vary in size and complexity. Table II. shows the main characteristics of the test functions used. It includes the name of the functions, the mathematical formula of each benchmark, the boundary defining the search space, and the function's dimensions. Categories for each benchmark are also provided: unimodal (U) and multimodal (M).

It should be noted that unimodal test functions have an optimization, while multimodal test functions have more than one optimization. Unimodal test functions are used to evaluate the ability to use optimization algorithms, while multi-mode test functions are used to evaluate the ability to explore optimization algorithms[23]. As shown in Table III, F1-F7 are single-state test functions, while F8-F23 are multi-state test functions. Also, the dimensions of the F14-F23 test functions are fixed.

In general, the improvement of the proposed method in obtaining the exact value of the optimal point and the evaluation of the function is shown in Table II based on the average values compared to the compared methods, which shows the efficiency of the proposed method in evaluating the optimization function. The accuracy of this method has not been reduced. The improved method is simulated in MATLAB, and the results of its comparison with its Gray Wolf Optimization (GWO)[4], Dragonfly Algorithm (DA)[24], and Bat Algorithm (BA)[25] are shown in Table II. As can be seen, in most cases, the proposed method performs better than the compared methods.

TABLE II

COMPARISON OF TEST FUNCTIONS BETWEEN SEVERAL OPTIMIZATION METHODS

Function	Criteria	BA	DA	GWO	BWO	OBBWO
F1	AVG	-2.28E-02	-2.37E-02	-2.37E-02	-2.37E-02	-4.00E-04
	BEST	-2.37E-02	-2.37E-02	-2.37E-02	-2.37E-02	-4.00E-04
	WORST	-2.26E-02	-2.37E-02	-2.37E-02	-2.37E-02	-4.00E-04
F2	AVG	8.02E-02	-2.37E-02	-2.37E-02	-2.37E-02	1.05E-01
	BEST	3.67E-02	-2.37E-02	-2.37E-02	-2.37E-02	8.13E-02
	WORST	1.18E-01	-2.37E-02	-2.37E-02	-2.37E-02	8.13E-02
F3	AVG	-2.17E-02	-2.37E-02	1.28E+00	1.21E-01	1.09E-01
	BEST	-2.22E-02	-2.37E-02	4.01E-01	7.65E-01	-1.95E-02
	WORST	-2.09E-02	-2.37E-02	4.71E+00	8.11E-01	-1.95E-02
F4	AVG	-1.04E-02	-2.37E-02	-9.03E-03	-2.36E-02	1.35E-01
	BEST	-2.11E-02	-2.37E-02	-2.31E-02	-2.37E-02	3.00E-03
	WORST	1.88E-02	-2.37E-02	7.17E-03	-2.26E-02	3.00E-03
F5	AVG	2.86E-01	9.11E+00	-1.54E-02	1.40E-01	1.03E+00
	BEST	-2.24E-02	-2.37E-02	-2.35E-02	-2.30E-02	8.65E-01
	WORST	5.23E-01	1.34E+02	5.23E-04	1.01E+00	8.65E-01
F6	AVG	-2.27E-02	-2.37E-02	-2.37E-02	-2.37E-02	1.03E+00
	BEST	-2.30E-02	-2.37E-02	-2.37E-02	-2.37E-02	-2.14E-02
	WORST	-2.25E-02	-2.37E-02	-2.37E-02	-2.37E-02	-2.14E-02
F7	AVG	-2.34E-02	-2.36E-02	3.67E-03	-2.07E-02	1.03E+00
	BEST	-2.37E-02	-2.37E-02	-9.24E-03	-2.18E-02	-2.04E-02
	WORST	-2.27E-02	-2.36E-02	1.28E-02	-1.77E-02	-2.04E-02
F8	AVG	-1.26E+04	-3.30E+03	-1.26E+04	-1.26E+04	-4.19E+03
	BEST	-1.26E+04	-3.74E+03	-1.26E+04	-1.26E+04	-4.19E+03
	WORST	-1.26E+04	-2.78E+03	-1.26E+04	-1.26E+04	-4.19E+03
F9	AVG	1.03E-02	1.14E+01	-2.37E-02	-2.37E-02	-4.19E+03
	BEST	-2.23E-02	3.96E+00	-2.37E-02	-2.37E-02	3.48E-01
	WORST	1.94E-01	1.89E+01	-2.37E-02	-2.37E-02	3.48E-01
F10	AVG	-1.21E-03	-2.37E-02	-2.37E-02	-2.37E-02	-4.19E+03
	BEST	-1.08E-02	-2.37E-02	-2.37E-02	-2.37E-02	4.43E-02
	WORST	2.62E-03	-2.37E-02	-2.37E-02	-2.37E-02	4.43E-02
F11	AVG	-1.99E-02	2.21E-01	-2.37E-02	-2.37E-02	-4.19E+03
	BEST	-2.37E-02	2.06E-02	-2.37E-02	-2.37E-02	7.25E-02
	WORST	-1.53E-03	5.52E-01	-2.37E-02	-2.37E-02	7.25E-02
F12	AVG	-2.37E-02	-2.37E-02	-2.37E-02	1.57E+32	-4.19E+03
	BEST	-2.37E-02	-2.37E-02	-2.37E-02	1.57E+32	-2.36E-02
	WORST	-2.37E-02	-2.37E-02	-2.37E-02	-2.37E-02	-2.36E-02
F13	AVG	-2.25E-02	-2.37E-02	-2.37E-02	-2.37E-02	-4.19E+03
	BEST	-2.36E-02	-2.37E-02	-2.37E-02	-2.37E-02	-2.34E-02
	WORST	-1.25E-02	-2.37E-02	-2.37E-02	-2.37E-02	-2.34E-02
F14	AVG	9.74E-01	9.74E-01	9.74E-01	9.74E-01	-4.19E+03
	BEST	9.74E-01	9.74E-01	9.74E-01	9.74E-01	9.74E-01
	WORST	9.74E-01	9.74E-01	9.74E-01	9.74E-01	9.74E-01
F15	AVG	-2.34E-02	-2.31E-02	-2.34E-02	-2.32E-02	-4.19E+03
	BEST	-2.34E-02	-2.34E-02	-2.34E-02	-2.34E-02	-2.32E-02
	WORST	-2.34E-02	-2.25E-02	-2.33E-02	-2.29E-02	-2.32E-02
F16	AVG	-1.06E+00	-1.06E+00	-1.06E+00	-1.06E+00	-4.19E+03
	BEST	-1.06E+00	-1.06E+00	-1.06E+00	-1.06E+00	-1.05E+00
	WORST	-1.06E+00	-1.06E+00	-1.06E+00	-1.06E+00	-1.05E+00
F17	AVG	3.74E-01	3.74E-01	3.74E-01	3.98E+01	-4.19E+03
	BEST	3.74E-01	3.74E-01	3.74E-01	3.98E+01	3.74E-01
	WORST	3.74E-01	3.74E-01	3.74E-01	3.74E-01	3.74E-01
F18	AVG	2.98E+00	2.98E+00	2.98E+00	2.98E+00	-4.18E+03
	BEST	2.98E+00	2.98E+00	2.98E+00	2.98E+00	2.98E+00
	WORST	2.98E+00	2.98E+00	2.98E+00	2.98E+00	2.98E+00
F19	AVG	-3.89E+00	-3.89E+00	-3.89E+00	-3.89E+00	-4.19E+03
	BEST	-3.89E+00	-3.89E+00	-3.89E+00	-3.89E+00	-3.88E+00
	WORST	-3.89E+00	-3.89E+00	-3.89E+00	-3.89E+00	-3.88E+00
F20	AVG	-3.31E+00	-3.24E+00	-3.35E+00	-3.35E+00	-4.19E+03
	BEST	-3.35E+00	-3.35E+00	-3.35E+00	-3.35E+00	-3.33E+00
	WORST	-3.23E+00	-3.23E+00	-3.35E+00	-3.35E+00	-3.33E+00
F21	AVG	-1.02E+01	-1.02E+01	-1.02E+01	-1.02E+01	-4.20E+03
	BEST	-1.02E+01	-1.02E+01	-1.02E+01	-1.02E+01	-1.01E+01
	WORST	-1.02E+01	-1.02E+01	-1.02E+01	-1.02E+01	-1.01E+01
F22	AVG	-1.02E+01	-1.04E+01	-1.04E+01	-1.04E+01	-4.21E+03
	BEST	-1.02E+01	-1.04E+01	-1.04E+01	-1.04E+01	-1.04E+01
	WORST	-1.02E+01	-1.04E+01	-1.04E+01	-1.04E+01	-1.04E+01
F23	AVG	-1.02E+01	-1.00E+01	-1.06E+01	-1.06E+01	-4.22E+03
	BEST	-1.02E+01	-1.06E+01	-1.06E+01	-1.06E+01	-1.05E+01
	WORST	-1.02E+01	-5.20E+00	-1.06E+01	-1.06E+01	-1.05E+01

TABLE III

FEATURES OF 23 BENCHMARKS. (N: DIMENSION, C: CATEGORY U: UNIMODAL, M: MULTIMODAL)

Name Function	Function	n	Range	C	n
Sphere	$f_1(x) = \sum_{i=1}^{n} x_i^2$	30	[-100, 100]	U	30
Schwefel's 2.22	$f_2(x) = \sum_{i=1}^{n} \lvert x_i^2 \rvert + \prod_{i=1}^{n} \lvert x_i^2 \rvert$	30	[-100, 100]	U	30
Schwefel's 1.20	$f_3(x) = \sum_{i=1}^{n} (\sum_{j=1}^{i} x_j)^2$	30	[-100, 100]	U	30
Schwefel's 2.21	$f_4(x) = \max\{\lvert x_i \rvert, 1 \leq i \leq n\}$	30	[-100, 100]	U	30
Rosenbrock	$f_5(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2]$	30	[-30,30]	U	30
Step	$f_6(x) = \sum_{i=1}^{n} ([x_i + 0.5])^2$	30	[-100,100]	U	30
Quartic Noise	$f_7(x) = \sum_{i=1}^{n} i x_i^4 + random[0,1)$	30	[-128,128]	U	30
Schwefel's 2.26	$f_8(x) = \sum_{i=1}^{n} -x_i \sin(\sqrt{\lvert x_i \rvert})$	30	[-500, 500]	M	30
Rastrigin	$f_9(x) = \sum_{i=1}^{n} [(x_i^2 - 10\cos(2\pi x_i) + 10)]$	30	[-5.12, 5.12]	M	30
Ackley	$f_{10}(x) = -20 \exp\left(-0.2 \sqrt{\frac{1}{n} \sum_{i=1}^{n} x_i^2}\right) - \exp\left(\frac{1}{n} \sum_{i=1}^{n} \cos(2\pi x_i)\right) + 20 + \exp(1)$	30	[-32,32]	M	30
Griewank	$f_{11}(x) = \frac{1}{4000} \sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	30	[-600,600]	M	30
Pendlized	$f_{12}(x) = \frac{\pi}{n}\{10 \sin(\pi y_1) + \sum_{i=1}^{n-1} y_i - 1)^2[1 + 10\sin^2(\pi y_{1+1})] + (y_n - 1)^2\} + \sum_{i=1}^{n} u(x_i, 10,100,4) \quad y_i \quad 1$ $+ \frac{x_i + 1}{4} \quad u(x_i, a, k, m) \begin{cases} k(x_i - a)^2 & x_i > a \\ 0 & -a < x_i < a \\ k(x_i - a)^2 & x_i < -a \end{cases}$	30	[-50,50]	M	30
Generalized pendlized	$f_{13}(x) = 0.1\left\{\sin^2(3\pi x_1) + \sum_{i=1}^{n} (x_i - 1)^2[1 + \sin^2(3\pi x_i + 1)] + (x_i - 1)^2[1 + \sin^2(2\pi x_n + 1)]\right\}$ $+ \sum_{i=1}^{n} u(x_i, 5,100,4)$	30	[-50,50]	M	30
Foxholes	$f_{14}(x) = (\frac{1}{500} + \sum_{j=1}^{25} \frac{1}{j + \sum_{i=1}^{2} (x_i - a_i)^6})^{-1}$	2	[-65,53]	M	2
Kowalik	$f_{15}(x) = \sum_{i=1}^{11} ([a_i - \frac{x_i(b_i^2 + b_i x_2)}{b_i^2 + b_i x_3 + x_4}])^2$	4	[-5,5]	M	4
Six-hump camel back	$f_{16}(x) = 4x_1^2 - 2.1x_1^4 + \frac{1}{3} x_1^6 + x_1 x_2 - 4x_2^2 + 4x_2^4$	2	[-5,5]	M	2
Branin RCOS	$f_{17}(x) = (x_2 - \frac{5.1}{4\pi^2} x_1^2 + \frac{5}{\pi} - 6)^2 + 10\left(1 - \frac{1}{8\pi}\right) \cos x_1 + 10$	2	[-5,5]	M	2
Goldstein Price	$f_{18}(x) = [1 + (x_1 + x_2 + 1)^2(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1 x_2 + 3x_2^2)] \times [30 + (2x_1 - 3x_2)^2$ $\times (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1 x_2 + 27x_2^2)]$	2	[-2,2]	M	2
Hartman 3	$f_{19}(x) = -\sum_{i=1}^{4} c_i \exp(-\sum_{j=1}^{3} a_{ij} (x_j - p_{ij})^2)$	3	[1,3]	M	3
Hartman 6	$f_{20}(x) = -\sum_{i=1}^{4} c_i \exp(-\sum_{j=1}^{6} a_{ij} (x_j - p_{ij})^2)$	6	[0,1]	M	6
Shekel 5	$f_{21}(x) = -\sum_{i=1}^{5} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	[0,10]	M	4
Shekel 7	$f_{22}(x) = -\sum_{i=1}^{7} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	[0,10]	M	4
Shekel 10	$f_{23}(x) = -\sum_{i=1}^{10} [(X - a_i)(X - a_i)^T + c_i]^{-1}$	4	[0,10]	M	4

VI. PRACTICAL EXAMPLES FOR THE PROPOSED METHOD

Here is a practical example of the proposed method. This algorithm has been used in feature selection to have a practical purpose for this algorithm. Psychological analysis of the text is challenging for the computer. The main reason for this is the inability of the computer to understand the user. Another purpose of this paper is to design a psychological analysis system of texts and a new way to improve it. So far, different features have been used in this field. In this paper, the BWO method, opposition-based and multi-objective, selects the feature, a new practical example for the proposed method. The feature selection steps are shown below, and its flowchart is shown in Figure 1, respectively. It first receives several texts and extracts its features and words as general features, which may be huge. The next step removes redundant letters and words and does not affect achieving the goal of the text's psychological analysis, such as prepositions, question words, auxiliary verbs, definite letters, etc.[26]. This step is called the preprocessing of texts. The next step is to use a machine learning method by giving texts and the objective function. Calculating the cost function uses a two-layer neural network, the hidden layer of sigmoid and ten neurons, the second layer is a linear layer and one neuron, and the Levenberg-Marquardt algorithm is used for this training network. This network takes inputs and targets, then creates and trains the neural network, and finally returns the results. It can be used to create a cost function.

The cost function is a two-objective function: the error percentage, and the other is the number of features. In the next step of multi-targeting the OBBWO method, several properties are selected based on two criteria: error rate and the number of features. The last step is to use the simplest classification method that most researchers [7, 27-30] used, namely the KNN interest classifier. Therefore, in the proposed method, we used the KNN classifier to evaluate the proposed algorithm's features and other algorithms more accurately. After performing these steps, the results are shown in Table V. The proposed method consists of five main steps for classification, which are shown in Figure 1. The evaluation criteria are defined in terms of the dissertation's four variables: TP, FP, TN, FN, the most important of which are listed below. Table IV shows the formula for each of these criteria[31].

TABLE IV
FORMULATION OF EACH CRITERION FOR EVALUATING THE PERFORMANCE OF CLASSIFIERS UNDER SUPERVISION

Criterion	Equation
accuracy	$\frac{tp + tn}{N}$
sensitivity(recall)	$\frac{tp}{tp + fn}$
specificity	$\frac{tn}{tn + fp}$
precision	$\frac{tp}{tp + fp}$
F-MEASURE	$f = 2\times\frac{precision\times recall}{precision+recall}$

Accuracy, one of the most important parameters for evaluating a classifier, is defined by monitoring and indicating the classifier's accuracy.

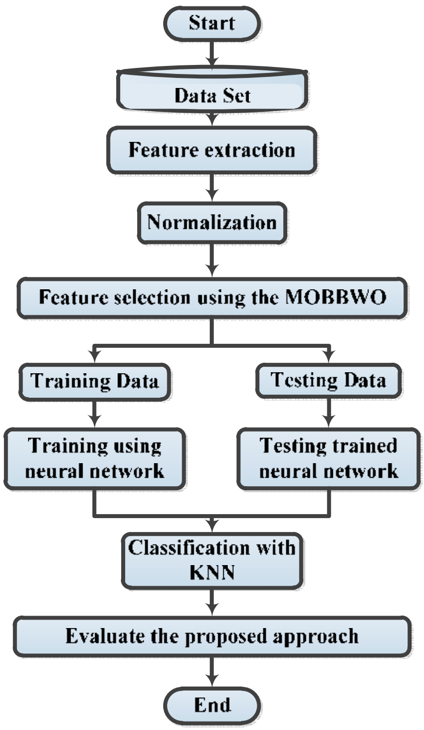

Fig.1. Applied example architecture

The sensitivity or reminder, also called the tp rate, refers to the probability of positive diagnoses from actual diagnoses. Specificity, also known as the tn rate, is probably due to correct negative diagnoses. Accuracy, also called a positive prediction rate, refers to the percentage of relevant predictions identified as relevant. F-measure is a combination of accuracy and reminder, and the closer to 1, the better.

1. Evaluation of practical example

The proposed algorithm in the MATLAB R2014a is simulated with a computer with a 64-bit i5 CPU and 4G memory. For this purpose, two sets of data called ISEAR and Sentiment polarity datasets v2.0 have been used. We use 80% of the data for training and the remaining 20% for testing. Table 5 shows the classifier performance of the proposed text with the existing basic algorithms. It should be noted that some of the algorithms implemented in MATLAB have been used for comparable algorithms.

Table V, it can be seen that the proposed algorithm, the improved BWO algorithm, has several better performance objectives in feature selection. The following table VI shows the proposed algorithm and other algorithms' results according to the average feature selections.

TABLE V
COMPARISON OF THE PERFORMANCE OF BASIC TEXT CLASSIFICATION ALGORITHMS WITH THE PROPOSED ALGORITHM

Dataset	Feature Selection	Number of Features	Accuracy	Sensitivity	Specificity	Precision	f-measure
ISEAR	MOBBWO	20	**0.933**	0.8312	0.95	0.8979	0.8354
		30	**0.9052**	0.9163	0.9693	0.9052	0.9352
		40	**0.9812**	0.9359	0.9313	0.9385	0.9379
	MIFS	20	0.9116	0.7192	0.9162	0.8685	0.7085
		30	0.9825	0.729	0.9552	0.7163	0.669
		40	0.9092	0.8985	0.9712	0.9385	0.9159
Sentiment polarity datasets v2.0	MOBBWO	100	**0.9412**	0.8309	0.8985	0.725	0.7359
		1000	**0.9616**	0.825	0.9637	0.7985	0.7987
		10000	**0.9383**	0.8491	0.938	0.8072	0.807
	MIFS	100	0.9314	0.5222	0.9671	0.6632	0.5658
		1000	0.9499	0.6376	0.9771	0.753	0.6761
		10000	0.9653	0.7915	0.9805	0.9581	0.8402

TABLE VI
COMPARISON OF THE PERFORMANCE OF BASIC TEXT CLASSIFICATION ALGORITHMS WITH THE PROPOSED ALGORITHM

DataSet	GA	BBOA	BSSA	BALO	BGWO	BDA	MOBBWO
ISEAR	**4**	5	6.5	8.2	4.5	4.3	**4**
Sentiment polarity datasets v2.0	15	14.5	13.1	20.8	13.2	10.3	**10.1**
Election	**4.6**	5.6	4.7	9.6	5.8	5.2	**5.3**
Healthcare	16	16	15.5	15	13	16.4	**11.2**
Sports	8	9.7	5.4	9	4.3	7	**4.1**

Table VI shows the MOBBWO approach results and other algorithms regarding the average number of feature selections that the MOBBWO approach has performed very well. So, this approach has proven its superiority in most data sets. Of course, it should be said that in the objective function, both the number of features and the accuracy of the classification are considered, so it is not possible to obtain a smaller number of features in the entire data set. The following compares the MOBBWO approach and other comparative algorithms in terms of classification accuracy, as shown in Table VII.

TABLE VII
RESULTS OF THE PROPOSED ALGORITHM AND OTHER ALGORITHMS IN TERMS OF OBJECTIVE FUNCTION

DataSet	GA	BBOA	BSSA	BALO	BGWO	BDA	MOBBWO
ISEAR	0.8378	0.5946	0.8784	0.6622	**0.9375**	0.8125	**0.9235**
Sentiment polarity datasets v2.0	0.7838	0.8649	0.726	0.6875	0.875	0.875	**0.9354**
Election	0.7264	0.7791	0.7937	0.7021	0.7998	0.7313	**0.8232**
Healthcare	0.8692	0.7733	0.8941	0.6344	0.8493	**0.9219**	**0.9121**
Sports	0.7447	0.6145	0.8140	0.6754	0.6185	0.7861	**0.8354**

Table VII shows the results of the MOBBWO approach and other algorithms in terms of classification accuracy. In addition to selecting the better feature shown in Table V, the MOBBWO approach also shows high performance in terms of classification accuracy. This approach has worked better in most datasets in terms of classification accuracy. These results are that in the objective function, both feature selection and categorization accuracy are considered. TableVIII shows the results of the proposed algorithm and other algorithms according to the average number of feature selections.

TABLE VIII
Results of the proposed two-way approach and other algorithms in terms of the average number of feature selections on the Sentiment analysis dataset

Iteration	GA	BBOA	BSSA	BALO	BGWO	BDA	MOBBWO
20	1941	2487	2381	2212	1833	1868	**1820**
40	1885	2120	1935	2145	1907	1887	**1814**
60	1850	1920	1896	2014	1754	2138	**1632**
80	1853	1865	1869	2352	1865	1821	**1810**
100	1815	1802	1803	1959	**1721**	2074	1754

Table VIII shows the proposed algorithm and other algorithms' results regarding the average number of feature selections that the MOBBWO approach has performed very well. This approach has proven itself in most repetitions.

VII. GENERAL EVALUATION AND CONCLUSION

In this paper, BWO is transformed into a discrete algorithm, then the same algorithm is based on contradiction, and finally, after multi-purpose, it is used in text analysis and psychology. The simulation results show that MOBBWO, as a new algorithm, scored 23 benchmarks and performed well in fifty different implementations. The average number of evaluations in these fifty implementations has been used to evaluate the degree of convergence and evaluate the optimization function. The proposed method results show that the improvement in the algorithm's convergence has occurred by reducing the number of evaluations of the optimization function, and this improvement has generally been more than twenty percent in the twenty-three benchmarks. The evaluation of the applied section in the text's psychological analysis also showed that the proposed method's classification accuracy is better than the compared methods and has a good performance in selecting features. In summary, the new method's advantages are: discretizing the BWO, converting the BWO to the conflict-based, multi-objective new method, using the choice of psychological features of Sentiment analysis.

REFERENCES

1. Gharehchopogh, F.S. and H. Gholizadeh, A comprehensive survey: Whale Optimization Algorithm and its applications. Swarm and Evolutionary Computation, 2019. 48: p. 1-24.
2. Gharehchopogh, F.S., I. Maleki, and Z.A. Dizaji, Chaotic vortex search algorithm: metaheuristic algorithm for feature selection. Evolutionary Intelligence, 2021: p. 1-32.
3. Abdollahzadeh, B. and F.S. Gharehchopogh, A multi-objective optimization algorithm for feature selection problems. Engineering with Computers, 2021: p. 1-19.
4. Jafari, N. and F. Soleimanian Gharehchopogh, An Improved Bat Algorithm with Grey Wolf Optimizer for Solving Continuous Optimization Problems. Journal of Advances in Computer Engineering and Technology, 2020. 6(3): p. 119-130.
5. Mohmmadzadeh, H. and F.S. Gharehchopogh, An efficient binary chaotic symbiotic organisms search algorithm approaches for feature selection problems. The Journal of Supercomputing, 2021: p. 1-43.
6. Rahnema, N. and F.S. Gharehchopogh, An improved artificial bee colony algorithm based on whale optimization algorithm for data clustering. Multimedia Tools and Applications, 2020. 79(43): p. 32169-32194.
7. Hosseinalipour, A., et al., A novel binary farmland fertility algorithm for feature selection in analysis of the text psychology. Applied Intelligence: p. 1-36.
8. 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.
9. Gharehchopogh, F.S., H. Shayanfar, and H. Gholizadeh, A comprehensive survey on symbiotic organisms search algorithms. Artificial Intelligence Review, 2019: p. 1-48.
10. 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.
11. Hosseinalipour, A., et al., Toward text psychology analysis using social spider optimization algorithm. Concurrency and Computation: Practice and Experience. n/a(n/a): p. e6325.
12. Dong, H., et al., A novel hybrid genetic algorithm with granular information for feature selection and optimization. Applied Soft Computing, 2018. 65: p. 33-46.
13. Liu, B. and L. Zhang, A survey of opinion mining and sentiment analysis, in Mining text data. 2012, Springer. p. 415-463.
14. Nasukawa, T. and J. Yi. Sentiment analysis: Capturing favorability using natural language processing. in Proceedings of the 2nd international conference on Knowledge capture. 2003.
15. Asghar, M.Z., et al., A review of feature extraction in sentiment analysis. Journal of Basic and Applied Scientific Research, 2014. 4(3): p. 181-186.
16. Saeys, Y., I. Inza, and P. Larrañaga, A review of feature selection techniques in bioinformatics. bioinformatics, 2007. 23(19): p. 2507-2517.
17. Sharma, M. and P. Kaur, A Comprehensive Analysis of Nature-Inspired Meta-Heuristic Techniques for Feature Selection Problem. Archives of Computational Methods in Engineering, 2020: p. 1-25.
18. Emine, B. and E. Ülker, An efficient binary social spider algorithm for feature selection problem. Expert Systems with Applications, 2020. 146: p. 113185.
19. Hayyolalam, V. and A.A.P. Kazem, BWO algorithm: A novel meta-heuristic approach for solving engineering optimization problems. Engineering Applications of Artificial Intelligence, 2020. 87: p. 103249.
20. Pang, B., L. Lee, and S. Vaithyanathan, Thumbs up? Sentiment classification using machine learning techniques. arXiv preprint cs/0205070, 2002.
21. Arora, S. and P. Anand, Binary butterfly optimization approaches for feature selection. Expert Systems with Applications, 2019. 116: p. 147-160.
22. Hussien, A.G., et al., S-shaped binary whale optimization algorithm for feature selection, in Recent trends in signal and image processing. 2019, Springer. p. 79-87.
23. Bennasar, M., Y. Hicks, and R. Setchi, Feature selection using joint mutual information maximisation. Expert Systems with Applications, 2015. 42(22): p. 8520-8532.
24. Mirjalili, S., Dragonfly algorithm: a new meta-heuristic optimization technique for solving single-objective, discrete, and multi-objective problems. Neural Computing and Applications, 2016. 27(4): p. 1053-1073.
25. Yang, X.S. and A.H. Gandomi, Bat algorithm: a novel approach for global engineering optimization. Engineering computations, 2012.
26. Leskovec, J., A. Rajaraman, and J.D. Ullman, Mining of massive data sets. 2020: Cambridge university press.
27. Mafarja, M.M. and S. Mirjalili, Hybrid whale optimization algorithm with simulated annealing for feature selection. Neurocomputing, 2017. 260: p. 302-312.
28. Liao, T.W. 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, 2018. 64: p. 581-595.
29. Mafarja, M., et al., Evolutionary population dynamics and grasshopper optimization approaches for feature selection problems. Knowledge-Based Systems, 2018. 145: p. 25-45.
30. Rajamohana, S. and K. Umamaheswari, Hybrid approach of improved binary particle swarm optimization and shuffled frog leaping for feature selection. Computers & Electrical Engineering, 2018. 67: p. 497-508.
31. Azar, A.T., et al., A random forest classifier for lymph diseases. Computer methods and programs in biomedicine, 2014. 113(2): p. 465-473.