Indonesian Journal of Electrical Engineering and Computer Science
Vol. 29, No. 2, February 2023, pp. 942~953
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v29.i2.pp942-953

Modified Harris Hawks optimizer for feature selection and support vector machine kernels

Hadeel Tariq Ibrahim[1], Wamidh Jalil Mazher[2], Enas Mahmood Jassim[3]
[1]E-Learning Unit, College of Education for Women, University of Thi-Qar, Thi-Qar, Iraq
[2]Department of Electrical, Al-Shatra Technical Institute, Southern Technical University, Basra, Iraq
[3]Department of Computer Science, College of Basic Education, University of Diyala, Diyala, Iraq

Article Info

Article history:

Received Jul 19, 2022
Revised Sep 28, 2022
Accepted Oct 18, 2022

Keywords:

Biomedical datasets
Feature selection
Harris Hawk optimizer
Heuristic optimization
Support vector machine

ABSTRACT

The support vector machine (SVM), one of the most effective learning algorithms, has many real-world applications. The kernel type and its parameters have a significant impact on the SVM algorithm's effectiveness and performance. In machine learning, choosing the feature subset is a crucial step, especially when working with high-dimensional data sets. These crucial criteria were treated independently in the majority of earlier studies. In this research, we suggest a hybrid strategy based on the Harris Hawk optimization (HHO) algorithm. HHO is one of the lately suggested metaheuristic algorithms that has been demonstrated to be used more efficiently in facing some optimization problems. The suggested method optimizes the SVM model parameters while also locating the optimal features subset. We ran the proposed approach HHO-SVM on real biomedical datasets with 17 types of cancer for Iraqi patients in 2010-2012. The experimental results demonstrate the supremacy of the proposed HHO-SVM in terms of three performance metrics: feature selection accuracy, runtime, and number of selected features. The suggested method is contrasted with four well-known algorithms for verification: firefly (FF) algorithm, genetic algorithm (GA), grasshopper optimization algorithm (GOA), and particle swarm algorithm (PSO). The implementation of the proposed HHO-SVM approach reveals 99.967% average accuracy.

Corresponding Author:

Hadeel Tariq Ibrahim
E-Learning Unit, College of Education for Women, University of Thi-Qar
64004 Thi-Qar, Iraq
Email: hadeel.tariq@utq.edu.iq

1. INTRODUCTION

Support vector machines (SVMs) are powerful tools in machine learning utilized to solve classification and regression problems [1]. The maturation of complex applications has made the employment of SVM vital [2]. SVM is a robust machine learning method for addressing classification and regression problems [3]. For the purpose of improving various cognitive and learning algorithms, bio-inspired systems have been thoroughly researched, SVM a popular supervised classification technique, is one of these algorithms. Vapnik was the one who initially devised and used SVM [4]. The SVM method attempts to find the ideal hyperplane that separates two classes by maximizing the distance between the edge of the hyperplane and the data points in the provided data set [5], [6]. One of the most well-known supervised models is the SVM algorithm, which is regarded as one of the best approaches in the field of machine learning. When compared to other techniques, SVM has certain strong advantages, including good generalization performance and the ability to produce high-quality decision limits founded on a trivial portion of training data points. Furthermore, the SVM excels in modelling intricate and non-linear relationships [7].

Different kernel functions have been employed by researchers to forecast SVM kernels. Because it changes just one parameter, the radial basis function (RBF) is a better function [8]. Cost (C) and gamma (γ), two SVM parameters, are modified by RBF [2]. SVM has been utilized in the literature for image retrieval [9], pattern recognition [10], human emotion recognition [11], spam categorization [12], cancer diagnoses [3], gender classification [13], and feature selection [14].

Despite the SVM algorithm's numerous benefits, it also had certain drawbacks, such as sensitivity to the parameter values at start-up. The cost (C) and kernel variables, like the gamma (γ) in the radial basis function (RBF) kernel, are among these variables. The simplification act of the SVM can be adversely affected by improper parameter selection. Along with this drawback, SVM is similar to many other machine learning algorithms in that its act is based on the features of the chosen data set, which is crucial for enhancing simplification performance, boosting computational efficiency, cutting down on running time, and producing very accurate classification models [15].

Harris Hawks optimizer (HHO) is a unique population-based, nature-inspired optimization algorithm. The cooperative attitude and surprise pounce pursuing technique of Harris' hawks in nature serve as the major sources of inspiration for HHO. In this clever tactic, many hawks work together to attack on a victim from various angles in an effort to surprise it. Founded on the dynamic nature of situations and the prey's fleeing movements, Harris Hawks can exhibit a variety of pursuit strategies. When compared to well-known metaheuristic methods, the HHO algorithm offers highly hopeful and irregularly competitive outcomes [16].

In this study, we present a brand-new HHO-SVM model that utilizes HHO in conjunction with SVM for the first time. This method makes use of HHO to concurrently perform feature selection and SVM parameter optimization. The model's objective is to use the fewest number of features while still maximizing SVM's classification accuracy. By comparing HHO-SVM with four other state-of-art algorithms, we demonstrated its high presentation. The other algorithms are the firefly (FF) algorithm [17], genetic algorithm (GA) [18], grasshopper optimization algorithm (GOA) [19], and particle swarm algorithm (PSO) [20].

The HHO algorithm works fast because it runs with a speed Levy and greedy choosing [21]. The proposed approach, HHO-SVM, is examined on (17) real biomedical datasets for Iraqi cancer patients [22], as listed Table 1. The proposed HHO-SVM results attained higher feature selection accuracy, lower runtime and fewer selected features compared to the other four algorithms.

The rest of this paper is structured as follows: the next section provides an outline of what has been done in the literature on some algorithms that have been employed in feature selection. Section 3 presents the basics of the Harris Hawk optimizer (HHO). The proposed HHO-SVM paradigm is discussed in section 4. In section 5, the experimental results are presented and analysed. Finally, in section 6, conclusions and future work are presented.

2. LITERATURE REVIEW

In feature selection, a variety of heuristic optimization strategies are used.; in this section, a fewheuristic optimization algorithms are presented. Huang and Wang [18] suggested and examined the usage of a genetic algorithm for instantaneously first choosing an optimum feature subset and second optimizing support vector regression factors (SVR) to increase the accuracy of the software power estimations. They described tests executed with two datasets of software plans. The simulations in both datasets showed that the suggested GA-based algorithm was capable of considerably improving the SVR performance. Khushaba *et al.* [23] modified differential evolution (DE) algorithm and proposed DEFS for feature selection. DEFS greatly decreased the computational costs and demonstrated robust performance. The DEFS approach was employed in a brain computer interface (BCI) application and compared with additional dimensionality lessening methods. Their results confirmed the importance of the proposed DEFS by obtaining an optimum solution and using less memory.

Lin *et al*. [24] developed particle swarm optimization (PSO) to determine the parameter and feature selection of the SVM, named PSO+SVM. They concurrently determined the support vector machine (SVM) kernel values while finding a feature subset without decreasing the accuracy of SVM classification [24]. The logistic and tent maps are two forms of chaotic maps that the particle swarm optimization (BPSO) technique depends on. In order to compute inertia, chaotic maps are employed as concealed in BPSO. In this approach, feature selection is highly accurate. The outcomes shown that the chaotic binary particle swarm optimization technique (CBPSO), which is based on the covering map, has greater accuracy than that of the logistic map [25].

The bat algorithm (BA), which is effectively utilized in feature selection, is modelled after how bats navigate flight pathways. BA doesn't need the usage of challenging operators like mutation and crossover. In essence, it alters the volume, frequency, and locations of bats. This approach guarantees accurate classification and lowers the size of the feature set [26]. By Emary *et al*. [17], the firefly algorithm (FF) was modified to propose a feature selection system. The modified FF was balanced adaptively to speed up the exploration and exploitation phases and find the optimum solution accordingly [17].

MVO in feature selection is based on employing a multiverse optimizer (MVO), a modern cosmology-inspired technique in selecting the best features and simultaneously optimizing the variables of the support vector machine (SVM). The outcomes shown that MVO can effectively reduce the number of characteristics picked while maintaining a high level of prediction accuracy [27].

By Emary *et al*. [28], a gray wolf optimizer was employed to find the optimum feature subset. In this paper, a comparison was performed with particle swarm optimization (PSO) and genetic algorithms (GAs) using a set of UCI data repositories. The authors approved the supremacy of the proposed algorithm in both classification accuracy and feature size minimization. Furthermore, the grey wolf optimization algorithm is more powerful than initialization in both PSO and GA optimizers.

The salp swarm algorithm [29] was developed to be used in feature selection. The accuracy and runtime of the proposed SSA-FS are compared with particle swarm optimization and differential evolution. In this study, bladder, breast, and colon cancers for Iraqi patients and synthetic datasets for evaluation were employed. The proposed SSA-FS attained the uppermost accuracies with shorter runtime compared with other selected algorithms. Ibrahim [19] optimized SVM parameters and selected features by a grasshopper optimization algorithm (GOA). It approved its capability to solve real-world issues with unknown search space.

3. HARRIS HAWK OPTIMIZER

The main approach of Harris hawks to hunt prey is “surprise pounce”, which is also known as the “seven kills” strategy. In this smart approach, some hawks go to supportively hit from diverse paths and concurrently converge on a perceived run away rabbit out the covering. This attack may speedily be done by arresting the surprised prey in limited seconds, but sometimes, concerning the run-away skills of the prey, the "seven kills" may consist of many short, fast rushes close to the prey in minutes [16].

3.1. Exploration phase

In HHO, Harris hawks lounge accidentally in some positions and wait to perceive a hunted rabbit founded on two strategies. The first strategy is modeled in (1) with considering an equal probability p for each lounging strategy, they lounge depending on the other family members' locations and the hunted animal (i.e., the rabbit) [16].

$$A(t+1) = \begin{cases} A_{rnd}(t) - rd_1 |A_{rnd}(t) - 2rd_2 A(t)| \; p \geq 0.5 \\ \left(A_{htd}(t) - A_{avg}(t)\right) - rd_3\left(L_{bnd} + rd_4(U_{bnd} - L_{bnd})\right) \; p < 0.5 \end{cases} \tag{1}$$

Where $A(t+1)$ represents the hawk position vector in the following iteration t, $A_{htd}(t)$ is the hunted rabbit location, $A(t)$ is the present hawk position vector, and rd_1, rd_2, rd_3, rd_4, and p are random numbers within (0,1) that are modified in every iteration. The upper and lower bounds of the parameters are represented by U_{bnd} and L_{bnd}, respectively. A randomly chosen hawk from the present population is denoted by $A_{rnd}(t)$, where A_{avg} represents the average location of the present hawk population. The average location of hawks is calculated by (2):

$$A_{avg}(t) = \frac{1}{M}\sum_{i=1}^{M} A_i(t) \tag{2}$$

where $A_i(t)$ represents each hawk location at iteration t and M indicates the entire number of hawks.

3.2. Exploration to exploitation transition

In the HHO algorithm, the transition from exploration to exploitation is done based on the prey escaping energy. The prey energy drops significantly through escape. In this step, the rabbit energy is demonstrated as:

$$P = 2P_0\left(1 - \frac{t}{T_{max}}\right) \tag{3}$$

where the rabbit run-out power is denoted by P, T_{max} is the maximum iteration, and the initial value of the rabbit power is denoted by P_0.

3.3. Exploitation phase

In this phase, Harris hawk birds achieve the "surprise pounce or seven kills" [30] by launching the purposed prey marked in the exploration phase. However, prey usually try to run in risky situations. Later, diverse hunting styles occurred in actual situations. As stated by the escape conduct of the prey and hunting strategies of Harris hawk birds, four probable strategies are suggested in the HHO algorithm to state the launching stage. By nature, prey always tend to run away from dangerous situations. The opportunity to run away is denoted by rd ; if the prey successfully runs away, $rd < 0.5$; otherwise, $rd \geq 0.5$.

4. THE PROPOSED HHO-SVM PARADIGM

The main goal of the proposed HHO-SVM is to select as few features as possible while maintaining increasing SVM classification accuracy. Here, not only collecting features in big datasets requires time and money but also redundant information consequences in wasting time during classification. Accordingly, it is better to lessen the number of features to obtain a quick response and to find a good relationship between the features and the results.

The implementation of three crucial components, including a search technique, an induction algorithm, and an assessment calculation, forms the basis of any wrapper feature selection approach [31]. In HHO-SVM, the HHO method is utilized as a search technique to find the best feature subset, and SVM is used as an induction algorithm, with assessments based on classification accuracy being used. Figure 1 displays the high-level structure of the wrapper feature selection together with a straightforward simulation of the method we suggested, known as the HHO-SVM.

The encoding characteristics and SVM parameters (i.e., C and γ), the goal function, and system architecture must all be taken into account while building the HHO-SVM paradigm. In the next subsections, these issues will be thoroughly explored.

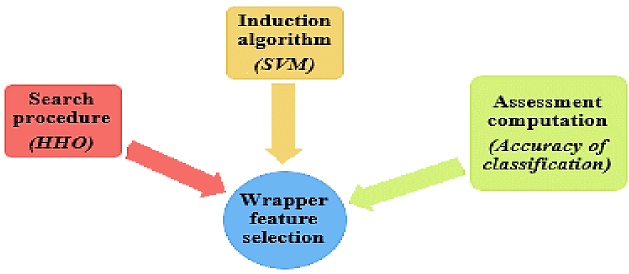

Figure 1. The elements of the proposed HHO-SVM algorithm's wrapper feature selection technique and their correspondences

4.1. Encoding SVM parameters and features

The first step of encoding is normalozing concurrently the inputted features and SVM parameters using (4) and (5) then the result is set in a vector. This vector comprises two portions: the first one contains SVM parameters (C, γ), where the second portion is for the selected features. First, SVM parameters are normalized, C to be in [0,4000] and γ in [0,30] interval using (4) [27].

$$Y = \frac{X - min_X}{max_X - min_X}(max_Y - min_Y) + min_Y \tag{4}$$

Where X and Y denoted to inputted C and γ respectively, $min_X = 0$, $max_X = 4000$, $min_Y = 0$, and $max_Y = 30$. Now, we apply (5) and then rounding features between [0,1]:

$$FB = \frac{FA - min_{FA}}{max_{FA} - min_{FA}} \tag{5}$$

where FA is the inputted feature, min_{FA} denoted to minimum value of it, max_{FA} is the maximum value. A feature is picked if the resulting FB value is larger than or equal to 0.5; otherwise, the value inside the vector is changed to 0, and no such feature is chosen.

4.2. Objective function

The objective function is needed in wrapper feature selection to assess the specific solution. The main aim of feature selection is to improve the accuracy of prediction and consequently minimize the number of selected features. In each selection of our proposed HHO-SVM system, the objective function is used based on calculation accuracy, as shown in (6) [32].

$$Accuracy = \frac{True_P + True_N}{True_P + False_N + False_P + True_N} \quad (6)$$

Where:
$True_P$: real class and all of the proper predictions are correct.
$True_N$: real class and all of the proper predictions are incorrect.
$False_N$: real class and all of the erroneous predictions are correct.
$False_P$: real class and all of the erroneous predictions are incorrect.

4.3. System layout

This section describes the layout of the proposed system, HHO-SVM, and lists its key components:
- Normalization of data: This feature selection approach involves public earlier processing. According to subsection 3.1, both SVM variables and features are normalized concurrently.
- Establishing training and testing sets: Each one of our biomedical datasets was split into a training set and testing set. The training set for the proposed HHO-SVM technique comprised 80% of the entire dataset, while the remaining 20% served as the testing set. We used the support vector machine (SVM) classifier to run the training and testing sets in order to create the model [33].
- Picking out a subset of features: Here, the value features for the 1 were selected from the training set.
- Assessment of fitness: The vectors from the designated training set have been utilized to control the classification act for SVM classifier learning, and (6).
- Breaking point: The top iteration has been determined, breaking the process altogether. The top iteration was really set to be at 5.

Figure 2 shows the planned HHO-SVM process and the relationships between the system's key components.

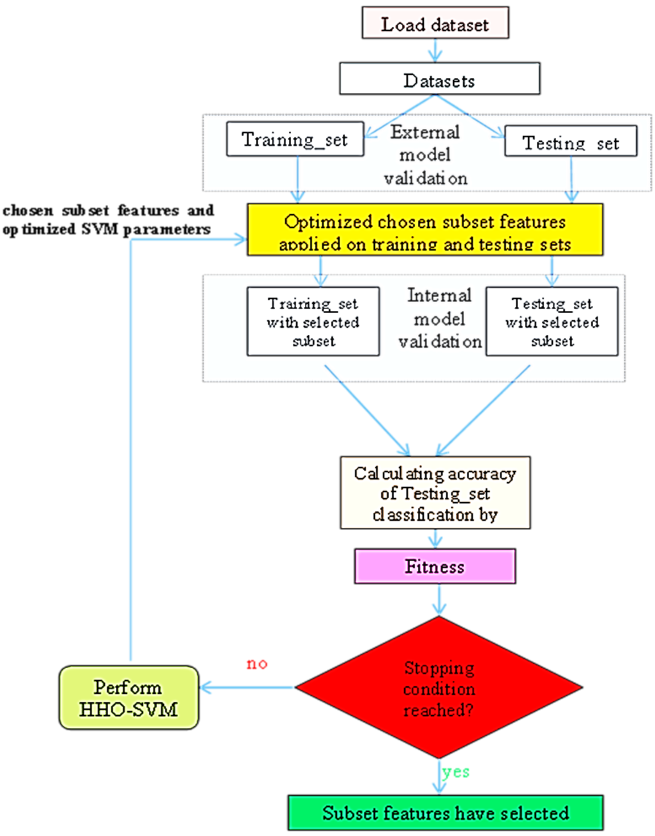

Figure 2. Proposed HHO-SVM workflow

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

In this study, we utilized 17 real datasets for differnt types of cancer in Iraq. The proposed HHO-SVM achieved the higher performance in most of the 17 real datasets. In addition, HHO-SVM is evaluated and contrasted with FF-SVM, GA-SVM, GOA-SVM, and PSO-SVM in terms of feature selection and SVM kernel factor optimization. These terms are:

- Feature selection accuracy.
- Run time (minutes: seconds: milliseconds).
- Number of selected features.

With an Intel(R) Core (TM) i7-5500U CPU running at 2.40 GHz, 8 GB of RAM, and Windows 10 as the operating system, we utilized MATLAB R2015a.

5.1. Datasets' explanation

Iraqi cancer patients' real biomedical datasets from 2010 to 2012 were used in this study [22]. For all cancer kinds, such datasets are gathered from all hospitals (public and private) in all Iraqi governorates. The final datasets included 16 features and various numbers of instances after being cleaned up of extraneous contains and bias values. Table 1 lists the specifics of the used datasets.

Table 1. List of datasets utilized in experiments

No	Dataset	No. of instances	No. of features
1	Abdomen	471	16
2	Bladder	4288	16
3	Blood	4788	16
4	Bones	950	16
5	Brain	2935	16
6	Breast	10670	16
7	Colon	3258	16
8	Eye	179	16
9	Glands	1655	16
10	Heart	183	16
11	Liver	2842	16
12	Lungs	4984	16
13	Lymph	5448	16
14	Naso	1818	16
15	Nerve	1175	16
16	Skin	1920	16
17	Stomach	2222	16

5.2. Comparisons of HHO-SVM with FF, GA, GOA, and PSO algorithms

5.2.1. Feature selection accuracy

The findings in Table 2 shows the comparisons of feature selection accuracy between HHO-SVM and the other four state-of-art algorithms with five iterations by each algorithm. Furthermore, the SVM classifier is employed in such a comparison without any optimization. Additionally, the optimized SVM parameters are listed in Table 2. Then, Table 2 accuracies are depicted by Figure 3. In 14 out of 17 datasets, HHO-SVM clearly outperformed other optimization algorithms in terms of feature selection accuracy (100%), as shown by the bold font. Consequently, as shown in Figure 4, HHO-SVM attained the greatest average accuracy of 99.967%. Moreover, GA excelled other algorithms on just three datasets whereas GOA obtained 100% over eight datasets.

The reason for this is that the progressive choice plan encourages search agents to modify their position over time and only pick the best options, allowing HHO to grow its concentration capabilities and solutions over the series of iterations with the maximum accuracy possible. GA sometimes quickly detects worthy solutions even for complex search spaces, and the procedure has some drawbacks associated with it. The main drawback is that the fitness function of the related problem should be well defined; otherwise, the GA may collide to local optima instead of the global-optimum solution [34]. This explains why the GA algorithm sometimes achieved high classification accuracies, but other times was not. The FF algorithm got the lowest accuracies because the FF algorithm needs an appropriate parameter setting with a numerous number of iterations to catch the optimum solution [35]. Due to the speedy convergence rate of PSO, it performs well and subsequently attains high accuracy [36].

Table 2. Comparison between proposed HHO-SVM and state-of-art algorithms based on classification accuracy in 5 iterations

Dataset		HHO-SVM	FF-SVM	GA-SVM	GOA-SVM	PSO-SVM	SVM (without Optimization)
Abdomen	Acc	100	81.528	99.954	91.549	99.921	92.958
	Cost (C)	981.322	3299.512	974.700	2528.256	480.199	-
	γ	0.084	0.016	0.002	0.2613	0.001	-
Bladder	Acc	100	87.523	99.956	100	99.887	90.278
	Cost (C)	882.214	2225.521	775.700	2422.257	310.200	-
	γ	0.084	0.032	0.003	0.270	0.003	-
Blood	Acc	99.746	86	99.909	99.653	99.867	92.014
	Cost (C)	782.524	4215.141	975.700	3528.168	670.200	-
	γ	0.009	0.022	0.014	0.005	0.002	-
Bones	Acc	100	73.894	99.934	78	99.865	88.667
	Cost (C)	2879.502	3116.742	1333.710	3125.711	8117.771	-
	γ	0.0214	0.025	0.001	0.002	0.001	-
Brain	Acc	100	75.604	99.949	100	99.833	82.222
	Cost (C)	992.213	3125.501	785.710	2551.207	221.201	-
	γ	0.033	0.022	0.023	0.281	0.005	-
Breast	Acc	100	69.736	99.970	100	99.926	85.294
	Cost (C)	882.211	4515.521	800.701	2422.257	311.201	-
	γ	0.084	0.032	0.023	0.271	0.053	-
Colon	Acc	99.613	82.473	99.954	99.612	99.866	94.574
	Cost (C)	2422.257	480.199	974.700	882.214	2422.257	-
	γ	0.001	0.021	0.008	0.101	0.051	-
Eye	Acc	100	70.391	99.966	89.655	99.939	86.207
	Cost (C)	311.201	3299.512	2422.257	775.700	311.201	-
	γ	0.014	0.018	0.022	0.015	0.004	-
Glands	Acc	100	82.356	99.970	87.097	99.933	87.097
	Cost (C)	974.700	882.214	2422.257	480.199	2422.257	-
	γ	0.271	0.281	0.311	0.001	0.282	-
Heart	Acc	100	78.688	99.961	93.939	99.956	93.939
	Cost (C)	981.322	2225.521	882.214	311.201	775.700	-
	γ	0.101	0.125	0.001	0.122	0.258	-
Liver	Acc	95	80.225	99.916	94.444	99.799	78.363
	Cost (C)	480.199	2422.257	670.200	2422.257	974.700	-
	γ	0.008	0.021	0.808	0.014	0.587	-
Lungs	Acc	100	89.626	99.941	100	99.843	76.823
	Cost (C)	3299.512	974.700	310.200	981.322	2422.257	-
	γ	0.272	0.001	0.205	0.311	0.288	-
Lymph	Acc	100	79.331	99.952	100	99.911	88.71
	Cost (C)	882.214	2422.257	2422.257	480.199	775.700	-
	γ	0.007	0.257	0.111	0.014	0.002	-
Naso	Acc	100	84.488	99.963	100	99.923	94.954
	Cost (C)	3299.512	2225.521	775.700	870.200	310.200	-
	γ	0.001	0.297	0.047	0.024	0.273	-
Nerve	Acc	100	74.893	99.969	95.429	99.932	96
	Cost (C)	775.700	2325.421	2422.257	2422.257	670.201	-
	γ	0.580	0.266	0.077	0.019	0.294	-
Skin	Acc	100	81.354	99.978	100	99.959	99.545
	Cost (C)	981.322	480.199	670.200	974.700	310.200	-
	γ	0.895	0.257	0.489	0.271	3325.523	-
Stomach	Acc	100	80.378	99.927	100	99.820	63.514
	Cost (C)	311.201	2225.521	2422.257	2335.541	8545.501	-
	γ	0.001	0.007	0.024	0.258	0.815	-
Average accuracy		99.967	79.911	99.951	95.846	99.893	87.715

5.2.2. Runtime

Obviously, runtime is extremely important to choose the right heuristic optimization algorithm, especially in higher dimensional search spaces [30]. Accordingly, in this study, we take into account calculating the runtime for all applied algorithms. As presented in Table 3, HHO-SVM confirmed its superiority to the FF-SVM, GA-SVM, GOA-SVM and PSO-SVM algorithms by consuming fewer runtimes over 8 datasets out of 17 datasets, as denoted by bold font. The minimum runtime has been achieved by HHO-SVM, as HHO performance is quick and competing in determining the right solutions [16]. In contrast, PSO outperformed the highest runtimes (as highlighted in Table 3) due its well-known stagnation ability into local optima, particularly in higher search space [36]. Accordingly, HHO-SVM achieved the lowest average runtime equal to 00:46:05 mm:ss:ms (minutes: seconds: millisecond), as shown in Figure 5. The proposed HHO-SVM is dominant from the runtime average view, where it consumes the least runtime average in comparison with the other four algorithms because the HHO algorithm runs with a fast Levy and greedy choosing [21].

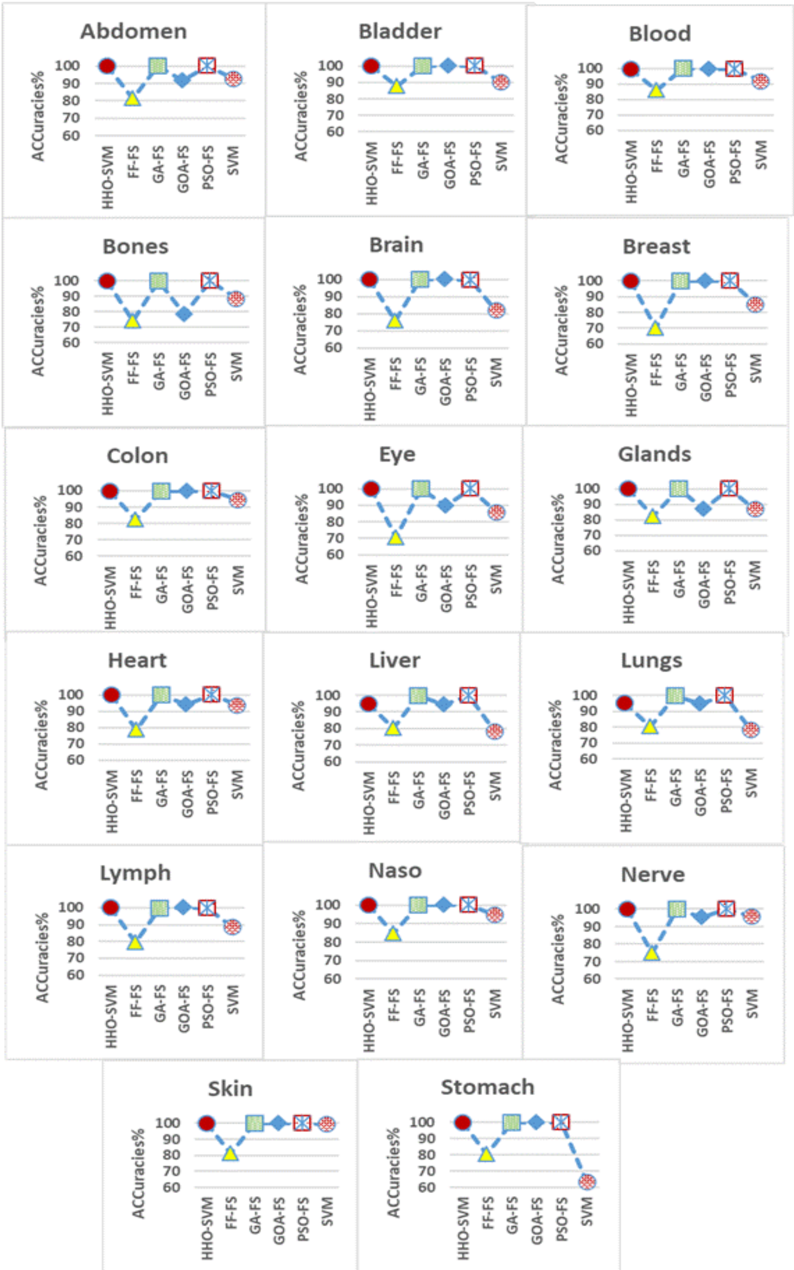

Figure 3. Comparison of feature selection accuracies between HHO-SVM and FF-SVM, GA-SVM, GOA-SVM, PSO-SVM, and SVM over 17 datasets

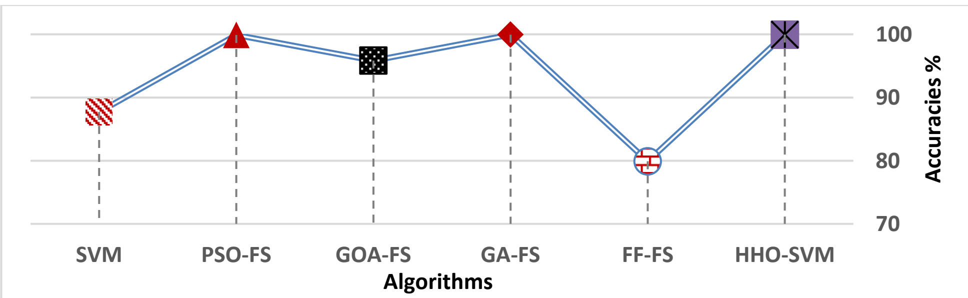

Figure 4. Comparison of feature selection average accuracies between HHO-SVM and FF-SVM, GA-SVM, GOA-SVM, PSO-SVM, and SVM over 17 datasets

Table 3. Comparison between proposed HHO-SVM and state-of-art algorithms based on runtime (mm:ss:ms)

Dataset	HHO-SVM	FF-SVM	GA-SVM	GOA-SVM	PSO-SVM
Abdomen	**00:02:39**	00:03:11	00:37:84	00:03:11	01:56:77
Bladder	**00:52:91**	01:56:16	03:31:85	01:10:66	12:26:35
Blood	01:28:86	03:05:91	03:03:75	**01:20:01**	11:01:50
Bones	00:08:60	00:08:24	00:50:65	**00:07:00**	02:19.71
Brain	00:54:19	01:10:98	02:13:87	**00:23:85**	07:29:58
Breast	**01:33:20**	03:53:71	04:02:41	04:56:99	16:04:90
Colon	00:49:57	01:19:33	02:34:27	**00:42:53**	05:27:97
Eye	**00:01:19**	00:01:48	00:30:48	00:01:51	01:25:59
Glands	**00:11:70**	00:20:89	01:08:09	00:13:68	03:15:31
Heart	00:01:01	**00:00:28**	00:47:96	00:01:33	01:29:46
Liver	01:04:03	00:58:05	02:31:32	**00:26:51**	05:44:36
Lungs	**01:11:74**	02:50:99	04:19:56	01:15:03	09:16:83
Lymph	03:25:56	03:53:24	**03:23:50**	03:34:60	17:12:40
Naso	**00:13:40**	00:26:21	01:35:91	00:18:19	05:47:91
Nerve	**00:08:25**	00:10:54	00:53:71	00:08:33	02:18:29
Skin	00:16:70	00:31:26	01:17:62	**00:15:31**	04:03:57
Stomach	00:35:21	00:41:85	01:40:34	**00:22:12**	04:46:58
Average	**00:46:05**	01:16:20	02:04:07	00:54:26	06:28:00

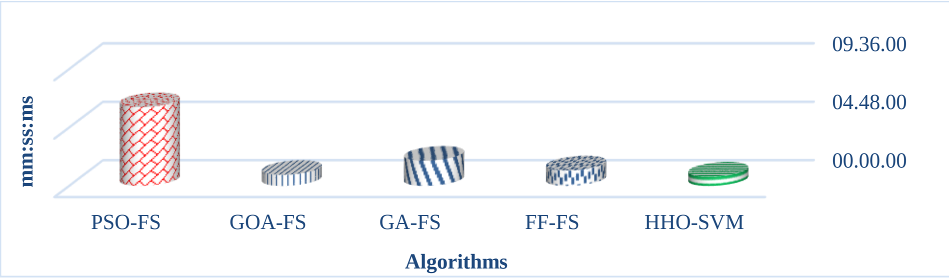

Figure 5. Comparison of runtime average accuracies between HHO-SVM and FF-SVM, GA-SVM, GOA-SVM, and PSO-SVM over 17 datasets

5.2.3. No. of selected features

In feature selection, the premium classifying algorithm must be able to outperform the smallest classification error rate by selecting the minimum number of features [37]. In Table 4 depicted with Figure 6, the minimum number of selected features is determined by the FF algorithm. FF outperformed the other algorithms on 10 datasets, and HHO-SVM outperformed the other algorithms on 8 datasets. As shown in Table 4, the FF and HHO-SVM algorithms achieved the lowest average of the selected features: 5.764 and 6, respectively. The comparison of selected features average between HHO-SVM and FF-SVM, GA-SVM, GOA-SVM, and PSO-SVM over 17 datasets is depicted in Figure 7.

Table 4. Comparison between proposed HHO-SVM and state-of-art algorithms based on number of selected features

Dataset	HHO-SVM	FF-SVM	GA-SVM	GOA-SVM	PSO-SVM
Abdomen	**4**	7	10	8	**4**
Bladder	**4**	5	8	9	10
Blood	6	6	**5**	11	11
Bones	8	**6**	9	10	10
Brain	**6**	**6**	7	8	7
Breast	8	**5**	**5**	12	9
Colon	7	**5**	6	6	9
Eye	6	**5**	8	11	10
Glands	**4**	7	6	9	8
Heart	7	7	8	10	9
Liver	**5**	**5**	8	7	11
Lungs	**5**	**5**	8	7	7
Lymph	7	6	11	9	10
Naso	**5**	6	7	6	8
Nerve	**5**	**5**	7	8	10
Skin	8	**6**	**6**	8	7
Stomach	7	**6**	8	9	6
Average	**6**	**5.764**	7.470	8.705	8.588

Obviously, HHO-SVM and GOA achieved higher accuracies, fewer runtimes, and nearly fewer selected averages. Finally, the minimum average of selected features is obtained by the FF algorithm. To assess the performances of the five mentioned algorithms, we must consider all three metrics. In other words, the victorious algorithm should outperform higher accuracy, less runtime and minimum number of selected features.

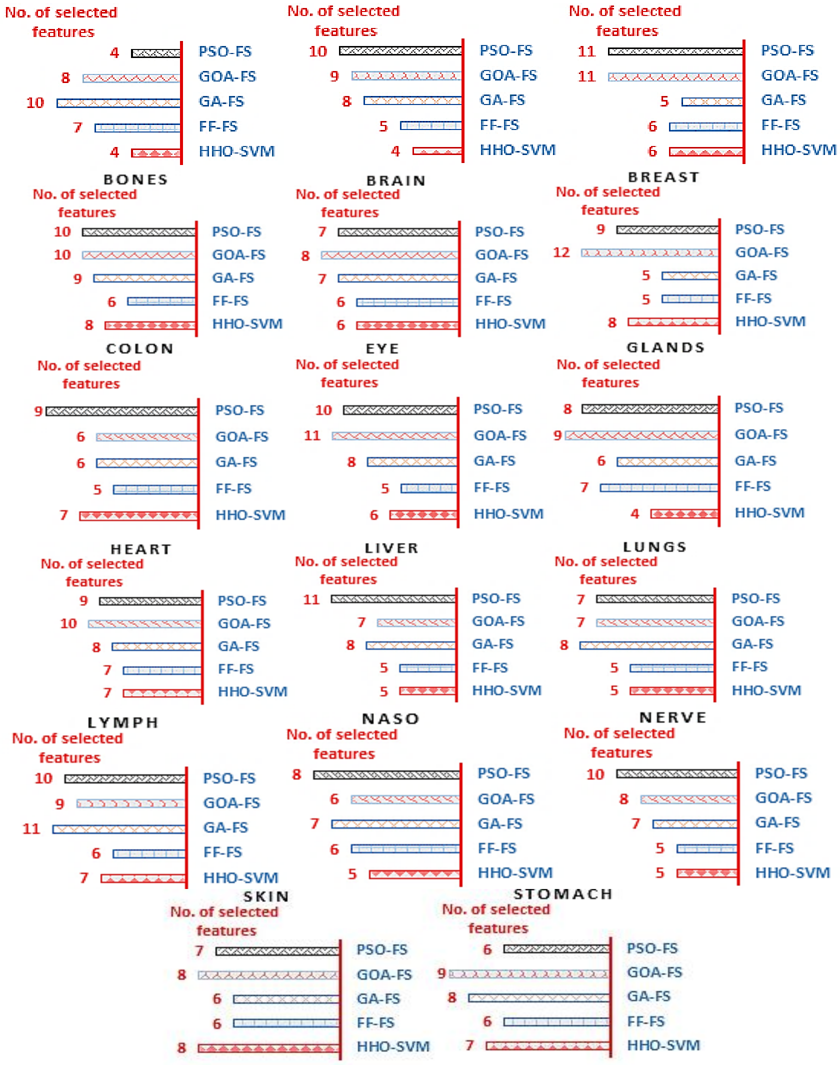

Figure 6. Comparison of no. of selected features between HHO-SVM and FF-SVM, GA-SVM, GOA-SVM, and PSO-SVM over 17 datasets

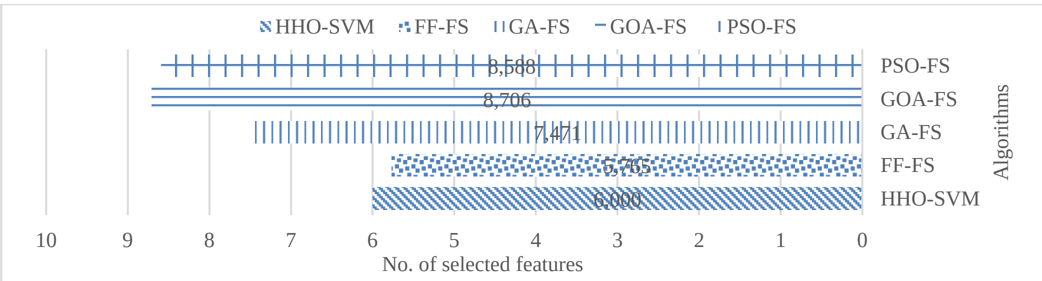

Figure 7. Comparison of selected features average between HHO-SVM and FF-SVM, GA-SVM, GOA-SVM, and PSO-SVM over 17 datasets

6. CONCLUSIONS

In this study, we provide a unique hybrid approach based on the Harris Hawk optimization algorithm (HHO) for SVM optimization. With the majority of the 17 actual datasets, the proposed HHO-SVM shown excellent performance. HHO-SVM approved its capability of finding the smallest and most effective subset of the model features while also adjusting the SVM kernel's parameters. This study demonstrates that improving SVM classifier performance by concurrently identifying the best kernel parameters and acceptable features improves classification-accuracy overall. Results of the experiments on the benchmark datasets demonstrated the HHO-SVM efficacy in improving the SVM classifier's accuracy. In most datasets, the HHO-SVM performs better in terms of classification accuracy than other optimizers including FF, GA, GOA, and PSO. Future research might look into and apply the suggested HHO-SVM model to other real-world word issues. Additionally, investigations of the model's performance on more complex issues are possible.

REFERENCES

[1] C. Staelin, “Parameter selection for support vector machines.” pp. 1–5, 2003, [Online]. Available: papers2://publication/uuid/F913CA32-08A3-432D-955E-A8F1EFAEAAE9.

[2] V. N. Vapnik, *The nature of statistical learning theory*. New York, NY: Springer New York, 2000.

[3] N. H. Sweilam, A. A. Tharwat, and N. K. A. Moniem, “Support vector machine for diagnosis cancer disease: a comparative study,” *Egyptian Informatics Journal*, vol. 11, no. 2, pp. 81–92, 2010, doi: 10.1016/j.eij.2010.10.005.

[4] V. Vapnik, “SVM method of estimating density, conditional probability, and conditional density,” *2000 IEEE International Symposium on Circuits and Systems. Emerging Technologies for the 21st Century. Proceedings (IEEE Cat No.00CH36353)*. Presses Polytech. Univ. Romandes, doi: 10.1109/iscas.2000.856437.

[5] W. Qiao and Z. Yang, “An improved dolphin swarm algorithm based on kernel fuzzy c-means in the application of solving the optimal problems of large-scale function,” *IEEE Access*, vol. 8, pp. 2073–2089, 2020, doi: 10.1109/access.2019.2958456.

[6] A. Gepperth and C. Karaoguz, “A bio-inspired incremental learning architecture for applied perceptual problems,” *Cognitive Computation*, vol. 8, no. 5, pp. 924–934, 2016, doi: 10.1007/s12559-016-9389-5.

[7] E. Tuba, L. Mrkela, and M. Tuba, “Support vector machine parameter tuning using firefly algorithm,” *2016 26th International Conference Radioelektronika (RADIOELEKTRONIKA)*. IEEE, 2016, doi: 10.1109/radioelek.2016.7477388.

[8] S. Salcedo-Sanz, J. L. Rojo-Álvarez, M. Martínez-Ramón, and G. Camps-Valls, “Support vector machines in engineering: an overview,” *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 4, no. 3, pp. 234–267, 2014, doi: 10.1002/widm.1125.

[9] L. Zhang, F. Lin, and B. Zhang, “Support vector machine learning for image retrieval,” *Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)*. IEEE, doi: 10.1109/icip.2001.958595.

[10] M. Lehtokangas, “Pattern recognition with novel support vector machine learning method,” *European Signal Processing Conference*, vol. 2015-March, no. March, 2000.

[11] N. A. A. Zulkifli, S. H. Sawal, S. A. Ahmad, and M. S. Islam, “Review on support vector machine (SVM) classifier for human emotion pattern recognition from EEG signals,” *Asian Journal of Information Technology*, vol. 14, no. 4, pp. 135–146, 2015, doi: 10.3923/ajit.2015.135.146.

[12] H. Drucker, D. Wu, and V. N. Vapnik, “Support vector machines for spam categorization,” *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 1048–1054, 1999, doi: 10.1109/72.788645.

[13] M.-H. Yang and B. Moghaddam, “Gender classification using support vector machines,” *Proceedings 2000 International Conference on Image Processing (Cat. No.00CH37101)*. IEEE, 2000, doi: 10.1109/icip.2000.899454.

[14] L. Hermes and J. M. Buhmann, “Feature selection for support vector machines,” *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. IEEE Comput. Soc, doi: 10.1109/icpr.2000.906174.

[15] J. Weston, S. Mukherjee, O. Chapelle, M. Pontil, T. Poggio, and V. Vapnik, “Feature selection for SVMs,” *Advances in Neural Information Processing Systems*, 2001.

[16] A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, “Harris hawks optimization: algorithm and applications,” *Future Generation Computer Systems*, vol. 97, pp. 849–872, 2019, doi: 10.1016/j.future.2019.02.028.

[17] E. Emary, H. M. Zawbaa, K. K. A. Ghany, A. E. Hassanien, and B. Parv, “Firefly optimization algorithm for feature selection,” *Proceedings of the 7th Balkan Conference on Informatics Conference*. ACM, 2015, doi: 10.1145/2801081.2801091.

[18] C.-L. Huang and C.-J. Wang, “A GA-based feature selection and parameters optimizationfor support vector machines,” *Expert Systems with Applications*, vol. 31, no. 2, pp. 231–240, 2006, doi: 10.1016/j.eswa.2005.09.024.

[19] H. T. Ibrahim, W. J. Mazher, O. N. Ucan, and O. Bayat, “A grasshopper optimizer approach for feature selection and optimizing SVM parameters utilizing real biomedical data sets,” *Neural Computing and Applications*, vol. 31, no. 10, pp. 5965–5974, 2018, doi: 10.1007/s00521-018-3414-4.

[20] L.-Y. Chuang, H.-W. Chang, C.-J. Tu, and C.-H. Yang, “Improved binary PSO for feature selection using gene expression data,” *Computational Biology and Chemistry*, vol. 32, no. 1, pp. 29–38, 2008, doi: 10.1016/j.compbiolchem.2007.09.005.

[21] S. Mirjalili, H. Faris, and I. Aljarah, “Introduction to evolutionary machine learning techniques,” *Algorithms for Intelligent Systems*. Springer Singapore, pp. 1–7, 2019, doi: 10.1007/978-981-32-9990-0_1.

[22] Ministry of Health-Iraq-Iraqi Cancer Board, “Acceptance of official cancer datasets from Iraq,” 2017. .

[23] R. N. Khushaba, A. Al-Ani, and A. Al-Jumaily, “Differential evolution based feature subset selection,” *2008 19th International Conference on Pattern Recognition*. IEEE, 2008, doi: 10.1109/icpr.2008.4761255.

[24] S.-W. Lin, K.-C. Ying, S.-C. Chen, and Z.-J. Lee, “Particle swarm optimization for parameter determination and feature selection of support vector machines,” *Expert Systems with Applications*, vol. 35, no. 4, pp. 1817–1824, 2008, doi: 10.1016/j.eswa.2007.08.088.

[25] C.-S. Yang, L.-Y. Chuang, J.-C. Li, and C.-H. Yang, “Chaotic maps in binary particle swarm optimization for feature selection,” *2008 IEEE Conference on Soft Computing in Industrial Applications*. IEEE, 2008, doi: 10.1109/smcia.2008.5045944.

[26] E. Emary, W. Yamany, and A. E. Hassanien, “New approach for feature selection based on rough set and bat algorithm,” *2014 9th International Conference on Computer Engineering & Systems (ICCES)*. IEEE, 2014, doi: 10.1109/icces.2014.7030984.

[27] H. Faris, M. A. Hassonah, A. M. Al-Zoubi, S. Mirjalili, and I. Aljarah, "A multi-verse optimizer approach for feature selection and optimizing SVM parameters based on a robust system architecture," *Neural Computing and Applications*, vol. 30, no. 8, pp. 2355–2369, 2017, doi: 10.1007/s00521-016-2818-2.

[28] E. Emary, H. M. Zawbaa, C. Grosan, and A. E. Hassenian, "Feature subset selection approach by gray-wolf optimization," *Advances in Intelligent Systems and Computing*. Springer International Publishing, pp. 1–13, 2015, doi: 10.1007/978-3-319-13572-4_1.

[29] H. T. Ibrahim, W. J. Mazher, O. N. Ucan, and O. Bayat, "Feature selection using salp swarm algorithm for real biomedical datasets," *IJCSNS International Journal of Computer Science and Network Security*, vol. 17, no. 12, 2017.

[30] J. A. Allen and S. Minton, "Selecting the right heuristic algorithm: runtime performance predictors," *Lecture Notes in Computer Science*. Springer Berlin Heidelberg, pp. 41–53, 1996, doi: 10.1007/3-540-61291-2_40.

[31] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 1–2, pp. 273–324, 1997, doi: 10.1016/s0004-3702(97)00043-x.

[32] M. F. Akay, "Support vector machines combined with feature selection for breast cancer diagnosis," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3240–3247, 2009, doi: 10.1016/j.eswa.2008.01.009.

[33] A. Mammone, M. Turchi, and N. Cristianini, "Support vector machines," *WIREs Computational Statistics*, vol. 1, no. 3, pp. 283–289, 2009, doi: 10.1002/wics.49.

[34] J. Guo, J. White, G. Wang, J. Li, and Y. Wang, "A genetic algorithm for optimized feature selection with resource constraints in software product lines," *Journal of Systems and Software*, vol. 84, no. 12, pp. 2208–2221, 2011, doi: 10.1016/j.jss.2011.06.026.

[35] L. Zhang, L. Liu, X.-S. Yang, and Y. Dai, "A novel hybrid firefly algorithm for global optimization," *PloS one*, vol. 11, no. 9, pp. e0163230–e0163230, Sep. 2016, doi: 10.1371/journal.pone.0163230.

[36] M. Li, W. Du, and F. Nian, "An adaptive particle swarm optimization algorithm based on directed weighted complex network," *Mathematical Problems in Engineering*, vol. 2014, pp. 1–7, 2014, doi: 10.1155/2014/434972.

[37] I. Aljarah *et al.*, "A dynamic locality multi-objective salp swarm algorithm for feature selection," *Computers & Industrial Engineering*, vol. 147, p. 106628, 2020, doi: 10.1016/j.cie.2020.106628.

BIOGRAPHIES OF AUTHORS

Hadeel Tariq Ibrahim is Assist Professor at College of Education for Women, University of Thi-Qar, Iraq. She Holds a PhD degree in Computer Science with specialization in software engineering. Her research areas are heuristic optimization algorithms, data mining, feature selection and machine learning. She is head of E-learning unit and Quality Assurance and Accreditation. She received the B.Sc. degree in computer science from the University of Baghdad, Iraq, the M.Sc. degree in Information Technology from Iraqi Commission for Computers and Informatics/Informatics Institute for Postgraduates Institute, Iraq, and the Ph.D. degree in Software Engineering from Altinbas University, Turkey. She can be contacted at email: hadeel.tariq@utq.edu.iq.

Wamidh Jalil Mazher is Assist Professor at Technical College, southern technical University, Iraq. He Holds a PhD degree in Communication Engineering with specialization in Optical communication. His research areas are communication, automation, support vector machine and machine learning. He is head of Electrical Dept. He received the B.Sc. degree in electrical engineering from the University of Technology, Iraq, the M.Sc. degree in computer and communication from UPM, Malaysia, and the Ph.D. degree in communication engineering from Altinbas University, Turkey. He can be contacted at email: wamidh.mazher@stu.edu.iq.

Enas Mahmood Jassim received the B.Sc. degree in computer science from University of Technology, Computer Science Dept., Iraq. She is programmer at Department of Computer Science, Basic Education College, University of Diyala, Iraq. She can be contacted at email: enasmahmoud@uodiyala.edu.iq.