Predicting Di-2-Ethylhexyl Phthalate Toxicity: Hybrid Integrated Harris Hawks Optimization With Support Vector Machines

BEIBEI SHI1, ALI ASGHAR HEIDARI2,3, CHENG CHEN4, MINGJING WANG5, CHANGCHENG HUANG6, HUILING CHEN6, (Associate Member, IEEE), AND JIAYIN ZHU7

1Zhenjiang First People’s Hospital, Affiliated People’s Hospital of Jiangsu University, Zhenjiang 212000, China
2School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran 1417466191, Iran
3Department of Computer Science, School of Computing, National University of Singapore, Singapore 117417
4Center of Clinical Research, The Affiliated Wuxi People’s Hospital of Nanjing Medical University, Wuxi 214023, China
5Institute of Research and Development, Duy Tan University, Da Nang 550000, Vietnam
6College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou 325035, China
7Laboratory Animal Center, Wenzhou Medical University, Wenzhou 325035, China

Corresponding authors: Changcheng Huang (cchuang@126.com) and Jiayin Zhu (zjiayin@wmu.edu.cn)

This work was supported in part by the Science and Technology Plan Project of Wenzhou, China, under Grant Y20190187, in part by the Talent Research Initiation Fund of Wenzhou Medical University under Grant QTJ19020, and in part by the Key P and D Program Projects in Zhejiang Province under Grant 2019C01041.

ABSTRACT Phthalic acid esters (PAEs) are organic pollutants and synthetic compounds and have adverse effects on human health. In this study, we investigated whether Di-2-Ethylhexyl phthalate (DEHP), one of many PAEs, has adverse effects on rats. Adult male Sprague-Dawley rats were treated daily by oral gavage with vehicle (corn oil) or DEHP at a dose of 3000 mg/kg/day for 15 days. The results showed that DEHP caused hepatotoxicity in rats. When compared with the control group, relative liver weights, and serum alanine aminotransferase levels significantly increased after DEHP exposure. Hepatocyte swelling and degeneration were also found in DEHP-exposed rats. This study proposes an effective intelligence framework for the prediction of DEHP poisoning. The framework is designed by integrating an enhanced Harris hawks optimization (HHO) with a support vector machine (SVM), which is called SGLHHO-SVM. The core characteristic of the developed methodology is the SGLHHO algorithm that integrates the levy mechanism and two core operators abstracted from the salp swarm algorithm and grey wolf optimizer to enhance and restore the search capabilities of the HHO. The presented SGLHHO approach is used to tackle the key parameter pair optimization of the SVM, and it is also utilized to grab the optimal feature subset. Regarding the optimal feature subset and the pair parameter simultaneously, SGLHHO-SVM can autonomously predict the DEHP poisoning. The developed SGLHHO was conducted on 23 benchmark problems and compared with other state-of-the-art and competitive methods. The results demonstrate that the designed SGLHHO performs superior to other competitors on most benchmark problems. Furthermore, the proposed SGLHHO-SVM is also compared with other machine learning algorithms on a real-life DEHP sampled data. Statistical results verify the proposal can show better predictive property and higher stability on all for metrics. Therefore, the SGLHHO-SVM may be served as a potential computer-aided tool for the prediction of DEHP poisoning.

INDEX TERMS Di-2-Ethylhexyl phthalate, hepatotoxicity, support vector machine, Harris hawks optimization, salp swarm algorithm, grey wolf optimizer.

I. INTRODUCTION Phthalic acid esters (PAEs) are commonly used as plasticizers which have been listed as emerging, xenobiotic, refractory, and hazardous pollutants. Di-2-Ethylhexyl phthalate (DEHP), one of many PAEs, has been listed as top priority pollutants by the United States Environmental Protection Agency and the European Union due to the great concern for DEHP that has been shown by toxicological studies in recent years. Green et al. [1] reported that total daily individual
ambient DEHP levels were approximately 0.27 mg per day in the USA. He et al. [2] reported that the concentrations and composition profiles of PAEs in biota, air, and water in an agricultural area of western China. It was found that DEHP concentrations are approximately 2.05-52.5 ng/g in fish, 1.00-9.09 ng/g in pig, 1.83-30.5 ng/g in chicken, 0.6-7.8 ng/g in cattle and 12.6-75.9 ng/g in vegetables. Moreover, in river water, indoor air, and outdoor air, DEHP is 73.3-261 ng/L, 1550-2300 pg/m3, and 713-2600 pg/m3, respectively. DEHP is easily transferred from plastic products to the environment, with annual usage in the world of more than 10 million tons, and it poses a threat to human health [3]. Exposure to DEHP has been found to be associated with liver, ovaries, mammary gland, testicular, and renal dysfunctions [4], [5]. Zhao et al. [6], [7] reported that DEHP exposure resulted in hepatotoxicity and renal injury in quail (Coturnix japonica) by reducing the expression of the heat shock response and antioxidant defense response. Therefore, the evaluation of DEHP exposure in organisms is still a critical issue.

Although some toxicological studies on DEHP have been done, there is little information on the detection and identification of DEHP exposure in rodents. The use of blood and biochemical measurements coupled with machine learning algorithms has emerged as an ideal approach for the identification of toxicity assessments [8], [9]. Williams et al. [10] developed a machine learning model with a Bayesian algorithm to predict drug-induced liver injury (DILI) and the accuracy up to 86%. He et al. [11] built a combined classifier to screen the hepatotoxic ingredients in Traditional Chinese Medicines by the NaiveBayes algorithm from eight machine learning algorithms (NaiveBayes, AdaboostM1, IBK, KStar, RandomForest, LibSVM, J48, and Bagging). They showed the accuracy was 72.8% and 78.3%, respectively [12]. Hammann et al. [13] present four machine learning models (k-nearest neighbor, artificial neural networks, decision tree induction, support vector machines) for the prediction of clinically relevant DILI from structure, and the accuracy up to 89%.

In this study, an augmented machine learning method is designed for the prediction of DEHP poisoning for the first time. In the developed methodology (SGLHHO-SVM), the levy mechanism and two core operators abstracted from the salp swarm algorithm and grey wolf optimizer strategy integrated harris hawks optimizer (SGLHHO) was designed to devise a reasonable support vector machine (SVM) for prediction of DEHP poisoning. First of all, the proposal SGLHHO algorithm was compared with several state-of-the-art competitors on 23 benchmark functions. The trial results verified the ascendant property of the SGLHHO approach. Secondly, the proposed SGLHHO-SVM was also conducted against some other metaheuristic algorithms (MEAs) and other methods for the prediction of DEHP poisoning. The motivation of this study is that design an augmented machine learning method for the prediction of DEHP poisoning, and integrating an enhanced Harris hawks optimization with a support vector machine as developed methodology. In this framework, levy mechanism and two core operators abstracted from the salp swarm algorithm and grey wolf optimizer are all integrated into original harris hawks optimization and then used to tackle the key parameter pair optimization of the SVM and grab the optimal feature subset simultaneously for autonomously predicting the DEHP poisoning.

The experimental results exhibited the apparent advantage of the proposed SGLHHO-SVM. The main contribution of this study is as below:

- An improved HHO, SGLHHO, is designed by integrating levy mechanism and two core operators abstracted from the salp swarm algorithm and grey wolf optimizer.
- The proposed SGLHHO has performed significant property on 23 function optimization tasks.
- The proposed SGLHHO successfully solved the parameter tuning for SVM for the first time.
- A potential methodology (SGLHHO-SVM) is treated as an efficient auxiliary tool for DEHP poisoning diagnosis in rats using only blood samples.

The paper is organized as bellow. A brief synopsis of the SVM, Levy mechanism, and two core operators is given in Section 2. Section 3 gives the proposed SGLHHO algorithm. The proposed SGLHHO-SVM model is shown in section 4. Section 5 describes the data collection and experimental setup. The simulation results of SGLHHO on benchmark functions and SGLHHO-SVM on the DEHP dataset are exhibited in Section 6. Section 7 gives a discussion on the results. The conclusion and future direction are exhibited in Section 8.

II. BACKGROUNDS

A. SUPPORT VECTOR MACHINE (SVM)

SVM is the most widely accepted machine learning based on risk minimization and VC dimensional theories. An excellent tradeoff must be achieved between minimization of the error of the training set and maximizing the margin to obtain the best generalization ability and avoid the phenomenon of overfitting. Owing to its satisfactory training speed and high classification accuracy, especially these small sample data sets, SVM has been applied in many cases [14]–[16].

The SVM can separate the pending samples with highest classification accuracy by hyperplane when a suitable w and b can be obtained. Regarding of kernel techniques, SVM can also solve nonlinear classification. The non-linear function can be modeled as bellow:

$$g(x) = sgn\left(\sum_{i=1}^{n} a_i y_i K(x^i, x) + b\right)$$  \hspace{1cm} (1)$$

where $K(x, x')$ is the kernel function and $K(x, x') = \text{exp}(-\gamma \|x - x\|^2)$ is Gaussian kernel. For more details, one can refer to [17].
B. LEVY MECHANISM AND TWO CORE OPERATORS

The levy mechanism [18] can be defined as follows:

\[ L(z) \sim z^{-1-\beta}, 0 < \beta \leq 2 \]  
\[ z = \frac{A}{|B|^\frac{1}{2\beta}} \cdot A \sim N(0, \sigma^2), B \sim N(0, \sigma^2) \]  
\[ \delta^2 = \left\{ \frac{\Gamma(1 + \beta)}{\beta \Gamma[(1 + \beta)/2]} \right\} \frac{\sin(\pi \beta/2)}{2^{\beta-1/2}} \]  

where \( z \) is the step size and \( \beta \) is the levy index, \( A/B \sim N(0, \sigma^2) \) means the samples originate from a Gaussian distribution, mean and variance are 0 and \( \delta^2 \), and \( \Gamma(\cdot) \) indicates the Gamma function. To further reduce the possibility of stagnation, the levy mechanism was adopted to update the positions of the agents, while the \( \beta \) is set to 3/2 in this study.

In addition, another two operators abstracted from salp swarm algorithm [19](SSA) and grey wolf optimizer [20] to further enhance and restore the search capability of the HHO. The core character of SSA is the mathematical model as follows:

\[ x_j^l = \begin{cases} F_j + c_1((a_j - b_j)c_2 + b_j), c_3 \geq 0 \\ F_j - c_1((a_j - b_j)c_2 + b_j), c_3 < 0 \end{cases} \]  

where \( x_j^l \) is the position of the leader salp in the \( j \)th dimension, \( F_j \) means the position of food source in the \( j \)th dimension, \( a_j \) and \( b_j \) are the upper bound and lower bound of \( j \)th dimension, \( c_1 \) is defines as equation 6, where \( l \) is the current iteration, \( L \) is the maximum number of iterations and \( c_2, c_3 \) are random number uniformly generated in \([0, 1] \).

\[ c_1 = 2e^{-(4l/L)^2} \]  

The position of the followers can be updated by equation 7, where \( i \geq 2, x_j^f \) means the position of ith follower salp in \( j \)th dimension, \( t \) is the time, \( v_0 \) is the initial speed, \( a = \frac{\text{var}}{v_0} \) and \( v = \frac{x - x_0}{t} \). For more details, one can refer to [19].

\[ x_j^f = \frac{1}{2}aT^2 + v_0t \]  

Another core operator is abstracted from the grey wolf optimizer, which is composed of social hierarchy, encircling prey, hunting, attacking prey, and search for prey parts. The core idea is the application of hierarchy. The main mathematical model is as follows:

\[ \bar{D}_a = |\bar{C}_1 \cdot \bar{X}_a - \bar{X}|, \bar{D}_b = |\bar{C}_2 \cdot \bar{X}_b - \bar{X}|, \bar{D}_c = |\bar{C}_3 \cdot \bar{X}_c - \bar{X}| \]  

where the \( \bar{D} \) means the distance between grey wolves and the prey, \( \bar{A} \) and \( \bar{C} \) are coefficient vectors by regarding of \( \bar{A} = 2\bar{a} \cdot \bar{r}_1 - \bar{a} \) and \( \bar{C} = 2 \cdot \bar{r}_2 \) respectively, \( \bar{X}_a, \bar{X}_b \) and \( \bar{X}_c \) are the positions of three closest wolves to current prey, the \( t \) indicates the current iteration. For more details, one can refer to [20].

\[ x_1 = \bar{X}_a - \bar{A}_1 \cdot (\bar{D}_a), \bar{X}_2 = \bar{X}_b - \bar{A}_2 \cdot (\bar{D}_b), \bar{X}_3 = \bar{X}_c - \bar{A}_3 \cdot (\bar{D}_c) \]  
\[ \bar{X}(t + 1) = \frac{\bar{X}_1 + \bar{X}_2 + \bar{X}_3}{3} \]  

III. PROPOSED SGLHHO

In a large number of scenarios, many different problems required achieving optimal or sub-optimal scheme to satisfy the decision-makers [21], [22]. A fact is that complexity and non-linearity of the practical problems generate computing challenges, which has resulted in various analytical models to be designed [23]–[27]. Due to the efficiency, metaheuristic algorithms (MEAs) can serve as one available approach to stochastically probe complex search terrain. As it is observed, a large of MEAs exists in the literature [28], [29]. Some well-known and creative methods are Genetic algorithm (GA) [30], particle swarm optimization (PSO) [31], [32], bacterial foraging optimization (BFO) [33], [34], teaching-learning based optimizer (TLBO) [35], [36], gray wolf optimizer (GWO) [20], [37], moth-flame optimization (MFO) [38], [39], grasshopper optimization algorithm (GOA) [40], [41], whale optimization algorithm (WOA) [42], [43], fruit fly optimization algorithm (FOA) [44], [45], slime mould algorithm (SMA) [46]. Due to its efficiency and effectiveness, these MEAs have been used in many problems, such as image segmentation in medical diagnosis, engineering design [47], feature selection [48], [49], workflow scheduling, and biopharmaceutical industry [39], [50]–[52]. As the latest member of MEAs, HHO performs a simple mechanism and can be used easily. Since proposed [53], it has been used widely [54]–[56], such as solar energy [57]–[59], feature selection [60], drug design and discovery [61]. Furthermore, a large number of improved HHO variants have been presented, for example, hybrid HHO-based sine cosine mechanism [62], Nelder-mead driven HHO [63], generalized Gaussian distribution HHO [64], multi-objective HHO [65], mutation strategies-based HHO [66], diversification enriched HHO [58], Multi-population version [67] random forest model based-HHO [68].

In this study, the levy mechanism and two core operators abstracted from the salp swarm algorithm and grey wolf optimizer have been integrated to enhance and restore the search capability of the HHO. As far as we know that it is the first time to introduce these two strategies into HHO. The procedure of the presented SGLHHO can be seen as follows.

In the first step, the initialization of the population \( x_i (i = 1, 2, 3, \cdots, N) \) with \( N \) agents are randomly generated within the established range \([lb, ub] \), then the fitness of each agent is calculated; finally, the agent who has minimum value is treated as the optimal solution (prey).

During the iteration, the candidate position of each agent in the population will be calculated according to its inherent mechanism of HHO, including exploration phase-based control parameter \(|E|\) and \( q \), and exploitation phase-based soft besiege and hard besiege, then if corresponding candidate position is better than its current one, then is updated by candidate one; otherwise, no change will be made.

Secondly, the core characteristic of SSA is performed by the population; The specific process can be shown: For the current population \( x_i (i = 1, 2, 3, \cdots, N) \) in the proposed
methodology, each $x_i$ will be calculated $F_j+c_1((ub_j-lb_j)c_2+lb_j)$ or $F_j-c_1((ub_j-lb_j)c_2+lb_j)$ decided by the value of random $c_3$, and then update $x_i$ regarding of the elite strategy. In addition, $c_1=2e^{-(4L/L)^2}$ can further increase the depth of particle search.

Thirdly, the core idea of hierarchy in the original GWO is also used to evolve the population, and the detailed can be seen in Figure 1. Through this processing mechanism, each agent in the population can learn from three leading wolves to promote the mechanism of the proposed algorithm by this \( \bar{X}(t+1) = \frac{\bar{X}_1+\bar{X}_2+\bar{X}_3}{3} \) equation. During this process, these three leading wolves are dynamically changing according to their respective fitness values.

Finally, the levy mechanism is adopted to adjust each agent in the population. When the position of ith agent $X(t+1)$ is updated, and the new candidate solution can be formulated as $X_i(t+1) = X(t+1) + levy(\beta)$, at the same the greedy strategy is also used to determine whether the current particle needs to be updated. This structure can effectively reduce the probability of the algorithm falling into local optimum. The detailed algorithm SGLHHO can be seen algorithm 1, and please refer to the original manuscript [53] for detailed formula explanation in this algorithm.

**IV. PROPOSED SGLHHO-SVM MODEL**
To predict DEHP poisoning, an improved SVM model based on SGLHHO (SGLHHO-SVM) was proposed in this study. The framework of SGLHHO-SVM is exhibited in Figure 2. In addition, the inner 5-fold and outer 10-fold cross-validation (CV) scheme are used simultaneously, which means the $K_1 = 5$ and $K_2 = 10$ in this framework. The core part of this model is SVM. In the input space, the radial basis function (RBF) kernel is used to map the aggregate data into a kind of hidden layer space. The process involves two key parameters (penalty coefficient $C$ and kernel width $\gamma$), and $n$ features subset, penalty parameter $C$ determines the tradeoff between fitting error minimization and model complexity and the kernel bandwidth $\gamma$ defines the non-linear mapping from the input space to some high-dimensional feature space. The SGLHHO algorithm evolves these two parameters and the optimal feature subset. In addition, the continuous space was converted through the sigmoid function into the binary space, and if the value is less than 0.5, the feature will be assumed to be chosen; otherwise, the features will be discarded. Finally, SVM makes an accurate diagnose of DEHP poisoning based on the obtained optimal $C$, $\gamma$, and feature subset.

**Algorithm 1 SGLHHO**

| Input | The population size $N$ and maximum number of iteration $T$ |
|-------|--------------------------------------------------|
| Output| The best position of the hawks $X_{\text{best}}$ and its fitness value $\text{fitness}_{\text{best}}$ |
| 1     | Initialize the random population $X_i(i=1,2,\ldots,N)$; |
| 2     | Calculate the fitness value of hawks; |
| 3     | Set the $X_{\text{rabbit}}$ as the location of rabbit (best position); |
| foreach $(X_i)$ do | |
| 6     | Update the initial energy $E_0 = 2\text{rand}() - 1$ and jump strength $J$; |
| 7     | Update the $E$ using $E = 2E_0(1-t/T)$; |
| 8     | if $|E| \geq 1$ then |
| 9     | if $q \leq 0.5$ then |
| 10    | Update the position using $X(t+1) = X_{\text{rand}}(t) - r_1[X_{\text{rand}}(t) - 2r_2X(t)]$ |
| 11    | if $q > 0.5$ then |
| 12    | Update the position using $X(t+1) = (X_{\text{rabbit}}(t) - \bar{X}_m(t)) - r_3(\text{lb} + r_4(\text{ub} - \text{lb}))$ |
| 13    | if $E < 1$ then |
| 14    | if $r > 0.5$, and $|E| \geq 0.5$ then |
| 15    | Update the position using $X(t+1) = \Delta X(t) - E|JX_{\text{rabbit}}(t) - X(t)|$ |
| 16    | if $r > 0.5$ and $|E| < 0.5$ then |
| 17    | Update the position using $X(t+1) = X_{\text{rabbit}}(t) - E|\Delta X(t)|$ |
| 18    | if $r < 0.5$, and $|E| \geq 0.5$ then |
| 19    | Soft besiege with progressive rapid dives; |
| 20    | if $r < 0.5$, and $|E| < 0.5$ then |
| 21    | Hard besiege with progressive rapid dives; |
| 22    | The population performs the core characteristic of SSA and the hierarchy in original GWO; |
| 23    | The levy mechanism is adopted to adjust the population; |
| 24    | return $X_{\text{best}}, \text{fitness}_{\text{best}}$ |
| 25    | End-Loop; |
V. EXPERIMENTAL DESIGNS

A. DATA COLLECTION

In this study, a total of 133 adult male SD rats weighing between \(220 \pm 20\) g were from the Laboratory Animal Centre (production license: SCXK(Zhe)2015-0001), Wenzhou Medical University, Wenzhou, China. Rats were kept in the animal house facility of the institute in standard polypropylene cages at \(23 \pm 3^\circ\)C temperature with 40-70% relative humidity, 12 h light-dark cycles (light: 07:00-19:00 h, dark: 19:00-07:00 h) with free access to ad libitum food and water. This work had received approval for research ethics from the Administration Committee of Experimental Animals, Laboratory Animal Center, Wenzhou Medical University, and a certificate of approval is available upon request with the inspection form number of wydw 2019-0448. All procedures were conducted in accordance with the guidelines of the Institutional Animal Care. Di-2-Ethylhexyl phthalate (DEHP, CAS no. 117-81-7, purity \(\geq 98\%\)) was obtained from Aladdin Industrial Corporation (Shanghai, China). All other solvents and reagents were obtained from commercial sources at the highest grade of purity.

In addition, histopathological examination was also conducted. Liver sections were collected from the experimental rats and immediately fixed in 4% formalin solution for 24 h, dehydrated in 70% ethanol. Liver tissues embedded in a paraffin block, and cut with a microtome to yield sections of 4 \(\mu m\) thickness for histopathological analysis. Sections were stained with hematoxylin-eosin (H&E) and blindly examined for the extent of liver damage under bright field microscope (Nikon Eclipse E200). Table 1 lists the detail of these features.

B. EXPERIMENTAL SETUP

Several other algorithms were involved as competitors, including EHHO [58], OBLHHO [69], HHO, ALO, MFO, WOA, GWO on a common benchmark function. The parameters of these algorithms are set up regarding the original papers, and the details can be seen in table 2. In the next experimental, the liver hypertrophy and unpaired Student t-test or one-way analysis of variance (ANOVA) (GraphPad Prism Software, San Diego, CA, USA, version 5.01) were both further investigated thoroughly. Then, the proposed algorithm SGLHHO was used to optimize the best parameter combination and feature subset for SVM, the resultant SGLHHO-SVM used for the prediction of DEHP poisoning on the collection data. SGLHHO-SVM was also compared with several popular machine learning methods, including original SVM, GWO-SVM, WOA-SVM, MFO-SVM, KELM, and KNN. The two key parameters including the \(C\) and \(\gamma\) SVM were set as \([-2^{-5}, 2^{-5}]\) and \([-2^{-5}, 2^{-5}]\) respectively. The data were scaled into the range [-1, 1] before the construction of the classifiers, to avoid the instability of the experiment caused by large data.

It should be noted that the simulation experiments were carried out in MATLAB environment ran on a Windows Server2016 R2 operating system with Intel (R) Xeon (R) CPU E5-2660 v3 (2.60 GHz) and 16GB of RAM. To make unbiased and objective results, the classification performance was assessed using the 10-fold cross-validation (CV) analysis. Furthermore, to assess the performance of the SGLHHO-SVM, four commonly used evaluation criteria, including specificity, sensitivity, classification accuracy (ACC), and Matthews Correlation Coefficients (MCC), were employed.
TABLE 1. List of the features used in this study and their definitions.

| Number | Features                        | Unit   |
|--------|---------------------------------|--------|
| X1     | total serum protein (TP)        | g/L    |
| X2     | gamma glutamyl transferase (GGT)| U/L    |
| X3     | alkaline phosphatase (ALP)      | U/L    |
| X4     | serum albumin (Alb)             | g/L    |
| X5     | albumin to globulin ratio (A/G) | /      |
| X6     | AST to ALT ratio (AST/ALT)      | /      |
| X7     | serum globulin (GLB)            | g/L    |
| X8     | serum alanine aminotransferase (ALT) | U/L          |
| X9     | serum aspartate aminotransferase (AST) | U/L          |
| X10    | serum urea nitrogen (Urea)      | mmol/L |
| X11    | serum uric acid (UA)            | umol/L |
| X12    | serum creatinine (Crea)         | umol/L |
| X13    | white blood cell (WBC)          | 109/L  |
| X14    | hemoglobin (HGB)                | g/L    |
| X15    | hematokrit (HCT)                | %      |
| X16    | mean corpuscular volume (MCV)   | fL     |
| X17    | mean corpuscular hemoglobin (MCH)|Pg        |
| X18    | mean corpuscular hemoglobin concentration (MCHC) | g/L    |
| X19    | blood platelet (PLT)            | 109/L  |
| X20    | lymphocyte percent (LYM%)       | %      |
| X21    | neutrophil percent (NEU%)       | %      |
| X22    | lymphocyte absolute count (LYM%)| /      |
| X23    | neutrophil absolute count (NEU%)| /      |
| X24    | platelet distribution width (PDW)|fL        |
| X25    | mean platelet volume (MPV)      | fL     |
| X26    | platelet-large cell ratio (P-LCR)|/      |
| X27    | eosinophil percent (EOS%)       | %      |
| X28    | monocytes absolute count (MON%) | 109/L  |
| X29    | monocytes percent (MON%)        | %      |
| X30    | eosinophil absolute count (EOS%)| 109/L  |
| X31    | thrombocytocrit (PCT)           | %      |
| X32    | red blood cell distribution width-standard deviation (RDW-SD) | fL     |
| X33    | red blood cell volume distribution width-coefficient variation (RDW-CV) | %      |
| X34    | red blood cell (RBC)            | 1012/L |

TABLE 2. Parameters setting for compared methods.

| Method   | Parameters                                                                 |
|----------|---------------------------------------------------------------------------|
| SGLHHO   | a=[2,0]; c=rand(); q=rand(); r=rand(); E0=2*rand()-1                     |
| EHHO     | q=rand(); r=rand(); E0=2*rand()-1                                         |
| OBLHHO   | q=rand(); r=rand(); E0=2*rand()-1                                         |
| HHO      | q=rand(); r=rand(); E0=2*rand()-1                                         |
| ALO      | /                                                                          |
| MFO      | b=1; a=[-1 1]; a ∈ [-1 - 2]                                               |
| WOA      | a1=[2 0]; a2=[-2 -1]; b=1                                               |
| GWO      | a=[2,0]                                                                   |

It can be defined as follows:

\[
ACC = \frac{TP + TN}{TP + FP + FN + TN} \times 100\% \quad (11)
\]

\[
Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (12)
\]

\[
Specificity = \frac{TN}{TP + FN} \times 100\% \quad (13)
\]

\[
MCC = \frac{A}{B} \quad (14)
\]

where TP and FN are the numbers of true positive cases and false negatives respectively. True negatives mean cases that are correctly categorized in the negative class. False positives are negative cases that are classified. TN and FP are the numbers of true negative and false positive. True negatives mean cases that are correctly categorized in the negative class. False positives are negative cases that are classified as positive, and \( A = (TP \times TN - FP \times FN) \), \( B = \sqrt{(TP + FP) \times (TP + FN) \times (TN + FP) \times (TN + FN)} \).

VI. EXPERIMENTAL RESULTS AND DISCUSSION

A. BENCHMARK FUNCTION VALIDATION

In this part, 23 benchmark functions were used to verify the performance of the SGLHHO, which is composed of 7 unimodal cases, 6 multimodal cases, and ten fixed dimensional multimodal cases. These benchmarks have been adopted in many works [70]–[72]. In addition, 30 independent tests were conducted to exclude the influence of random factors.

The presented SGLHHO in this study is compared with several algorithms, including EHHO [58], OBLHHO [69], HHO, ALO, MFO, WOA, GWO. All these methods were conducted on these 23 benchmarks by 30 independent executions. The detailed experimental results are exhibited in table 3, where the mean indicates the average of the results, and STD is the standard deviation. It can be seen that the average results of the presented SGLHHO are the lowest among these benchmarks, and it shows the apparent superiority. Furthermore, in terms of Wilcoxon's test, the SGLHHO is much
| Algorithm | F1 | F2 | F3 | F4 |
|-----------|----|----|----|----|
| SGLHOO    | 0.000E+000 | 0.000E+000 | 6.95E+216 | 1.71E+198 |
| BHHHO     | 1.71E-015 | 5.29E-015 | 3.74E+013 | 3.89E+013 |
| OBHHHO    | 2.34E+003 | 2.15E+003 | 4.41E+003 | 7.93E+000 |
| HHHO      | 3.75E-096 | 1.18E-095 | 1.53E-099 | 2.69E+003 |
| ALO       | 4.80E+001 | 7.22E+001 | 2.89E+001 | 1.48E+004 |
| MFO       | 4.95E+003 | 9.36E+003 | 4.38E+003 | 1.93E+007 |
| WOA       | 3.06E-044 | 9.65E-044 | 4.67E-035 | 8.53E+004 |
| GWO       | 4.96E-016 | 4.57E-016 | 4.64E-010 | 2.57E-001 |

**TABLE 3. Comparison results of SGLHHO with other peers on the benchmark functions.**

| Algorithm | F5 | F6 | F7 | F8 |
|-----------|----|----|----|----|
| SGLHOO    | 2.81E-001 | 2.22E+001 | 7.10E+000 | 1.00E+000 |
| BHHHO     | 2.95E+001 | 3.78E+001 | 5.66E+000 | 2.48E+000 |
| OBHHHO    | 0.000E+000 | 0.000E+000 | 1.74E+000 | 9.76E+000 |
| HHHO      | 2.86E+001 | 5.41E+001 | 3.75E+000 | 3.83E+000 |
| ALO       | 5.40E+001 | 9.82E+001 | 5.93E+001 | 7.74E+001 |
| MFO       | 1.54E+001 | 1.52E+001 | 3.27E+001 | 6.11E+001 |
| WOA       | 2.50E+000 | 3.70E+001 | 2.52E+000 | 5.14E+001 |
| GWO       | 2.82E-001 | 6.26E+001 | 1.98E+000 | 6.97E+000 |

| Algorithm | F9 | F10 | F11 | F12 |
|-----------|----|-----|-----|-----|
| SGLHOO    | 0.000E+000 | 0.000E+000 | 8.83E+000 | 1.00E+000 |
| BHHHO     | 0.000E+000 | 0.000E+000 | 2.84E+000 | 3.39E+000 |
| OBHHHO    | 0.000E+000 | 0.000E+000 | 2.84E+000 | 3.39E+000 |
| HHHO      | 0.000E+000 | 0.000E+000 | 2.84E+000 | 3.39E+000 |
| ALO       | 4.32E+001 | 3.31E+001 | 1.52E+000 | 8.30E+000 |
| MFO       | 1.91E+001 | 7.48E+001 | 1.92E+000 | 6.30E+001 |
| WOA       | 2.70E+001 | 4.37E+001 | 5.79E+001 | 4.76E+001 |
| GWO       | 9.76E+000 | 7.50E+001 | 6.68E+000 | 4.60E+001 |

| Algorithm | F13 | F14 | F15 | F16 |
|-----------|-----|-----|-----|-----|
| SGLHOO    | 2.79E+000 | 3.06E+002 | 9.00E+000 | 4.16E+000 |
| BHHHO     | 2.75E+000 | 7.65E+002 | 3.41E+000 | 2.81E+000 |
| OBHHHO    | 1.23E+000 | 1.32E+000 | 3.32E+000 | 2.69E+000 |
| HHHO      | 1.82E+000 | 1.26E+000 | 2.16E+000 | 3.68E+000 |
| ALO       | 2.39E+000 | 3.93E+000 | 4.91E+000 | 5.02E+000 |
| MFO       | 4.61E+000 | 8.35E+000 | 5.11E+000 | 3.96E+000 |
| WOA       | 1.68E+000 | 3.59E+000 | 5.34E+000 | 3.91E+000 |
| GWO       | 1.65E+000 | 2.14E+000 | 6.68E+000 | 3.99E+000 |

| Algorithm | F17 | F18 | F19 | F20 |
|-----------|-----|-----|-----|-----|
| SGLHOO    | 3.93E-001 | 2.54E+000 | 3.00E+000 | 1.93E+000 |
| BHHHO     | 4.09E-001 | 9.83E+000 | 3.00E+000 | 1.93E+000 |
| OBHHHO    | 4.28E-001 | 1.62E+000 | 3.00E+000 | 1.93E+000 |
| HHHO      | 3.97E-001 | 9.24E+000 | 3.00E+000 | 1.93E+000 |
| ALO       | 4.02E-001 | 5.33E+000 | 3.00E+000 | 1.93E+000 |
| MFO       | 3.97E-001 | 1.00E+000 | 3.12E+000 | 3.12E+000 |
| WOA       | 3.97E-001 | 3.23E+000 | 5.63E+000 | 2.32E+000 |
| GWO       | 3.96E-001 | 4.54E+000 | 3.10E+000 | 2.17E+000 |

| Algorithm | F21 | F22 | F23 |
|-----------|-----|-----|-----|
| SGLHOO    | -1.01E+001 | 8.75E+000 | -1.04E+001 |
| BHHHO     | -2.05E+000 | 1.84E+000 | -1.94E+000 |
| OBHHHO    | -2.83E+000 | 2.66E+000 | -3.19E+000 |
| HHHO      | -3.30E+000 | 1.05E+000 | -1.04E+000 |
| ALO       | -1.11E+001 | 7.22E+000 | -1.04E+000 |
| MFO       | -5.16E+001 | 3.53E+000 | -5.62E+000 |
| WOA       | -7.32E+000 | 2.49E+000 | -5.53E+000 |
| GWO       | -7.65E+000 | 3.31E+000 | -1.38E+000 |

**VOLUME 8, 2020**
better than other peers in 13 of 23 functions in a significant manner. Table 4 exhibits the sorting results of the Friedman test for SGLHHO versus all other competitors. According to the average ranking value of the involved algorithms, it can be seen that the SGLHHO shows the second-best performance on these benchmark functions, HHO, GWO, EHHO, WOA, OBLHHO, MFO, and ALO show the worse search performance. A reason can account for that the core characteristics of the involved SCA, GWO, and levy in this study can make original HHO obtain a better between exploration and exploitation.

To show the specific property of the SGLHHO, figure 3 records the convergence curves of these algorithms on some benchmark functions. It can be observed that the SGLHHO algorithm perform fast conference ability, and is superior to all other competitors in these functions. It should be noted that the SGLHHO can perform fast convergence search, especially F3, F7, which means it can get the theoretical optimal value quickly, but other competitors need more exploration and need to be accelerated. Furthermore, the same convergence trend can also be observed in F11, F15, F20, F21, F22, and F23. In sum, a conclusion can be made that these

|               | SGLHHO | EHHO  | OBLHHO | HHO   | ALO   | MFO   | WOA   | GWO   |
|---------------|--------|-------|--------|-------|-------|-------|-------|-------|
| +/-           | 16/3/4 | 19/3/1| 11/5/7 | 21/0/2| 18/2/3| 14/4/5| 15/5/3|       |
| ARV           | 2.6823 | 4.6514| 5.1506 | 2.3465| 6.1759| 5.6916| 4.7322| 3.821 |

**FIGURE 3.** Convergence curves selected benchmark functions.
B. Shi et al.: Predicting Di-2-Ethylhexyl Phthalate Toxicity: Hybrid Integrated Harris Hawks Optimization With SVMs

FIGURE 4. Relative weight of rat liver or kidneys for DEHP in rats. Data are presented as the mean ± SD. N = 7 rats per group. **Significantly different from the control group (P < 0.01).

FIGURE 5. DEHP-induced damage in hepatic tissue sections stained with H&E (x200). Rats were given corn oil (control) or DEHP (3000 mg/kg) orally once a day for 15 days. Hepatocyte swelling (arrows) and inflammatory infiltrates (arrowheads) were observed in DEHP exposure groups.

involved mechanisms can significantly enhance the property of the original HHO.

B. PREDICTION RESULTS OF DEHP POISONING

To investigate whether DEHP has toxic effects on rat liver, male adult rats were orally administrated 3000 mg/kg DEHP or corn oil as the control for 15 consecutive days. As shown in Figure 4, relative liver weights, the indicator of liver hypertrophy, were significantly increased in the DEHP exposure group compared with the control group. The results revealed that there was no difference in relative kidneys weights between DEHP exposure group and the normal control group. Figure 5 shows the transformation in collagen and tissue structure. There were pathological changes, including inflammation in the liver of the group that received DEHP. Hepatic tissues of DEHP treated rats revealed hydropic degeneration of hepatocytes, cell swelling. However, no significant changes were observed in the liver sections of the vehicle control rats. Histological considerations presented that the control group had a normal structural design of liver tissue as well as a normal monolayer of hepatocytes radially organized.

The blood parameters of the control group and the DEHP group were shown in Table 5. We observed that the levels of ALT and AST were significantly elevated in the DEHP exposure group when compared with the control group (P < 0.05). These results suggest that DEHP exposure induced rat liver injury.

In the next part, the effectiveness of the SGLHHO-SVM method for the prediction of DEHP poisoning was evaluated deeply. Table 6 shows the detailed results. It can be observed from this table that the classification accuracy fulfilled by the SGLHHO-SVM is 96.7%, Matthew’s correlation coefficient is 92.88%, sensitivity is 10%, specificity is 97.5%, and their variance is 0.0532, 0.1146, 0 and 0.0791 respectively. Besides, in this experiment, we can observe that the proposed SGLHHO-SVM can automatically obtain the optimal parameters of the SVM model, which is mainly due to the enhanced SGLHHO that can find the optimal parameters effectively.

To further verify the property of this model, the proposed methodology is compared with the other competitors SVM, HHO-SVM, GWO-SVM, WOA-SVM, MFO-SVM, KELM, and KNN. The comparison in terms of the four metrics and standard deviation is shown in figure 6. The results show that the SGLHHO-SVM algorithm shows superior to other competitors in four evaluation metrics, including ACC, MCC, sensitivity, and specificity, and its corresponding standard deviation is also the smallest among all models. This indicates that SGLHHO-SVM performs better performance and stability compared with the original model. In terms of the ACC evaluation metric, the SGLHHO-SVM model is the best. The results of the original SVM and original KELM perform the same property approximately. It is noted that the SVM-based original HHO has the second performance and just below the SGLHHO-SVM. The calculation results of the GWO-SVM and MFO-SVM are very close, and the KNN shows the worst; Regard to the MCC metric, SGLHHO-SVM still achieve the best results followed by the WOA-SVM, MFO-SVM, GWO-SVM, original SVM, KELM, and KNN. KNN still has the worst outcomes than other competitors and also the most substantial standard deviation among all these involved approaches. In terms of sensitivity
TABLE 5. Statistical analysis of 134 rats (control group and DEHP group).

| Indexes | Control(n=67) | DEHP (n=66) | P-Value |
|---------|---------------|-------------|---------|
| Tp      | 47            | 74.6        | 0.008   |
| GGT     | 0             | 8.1         | 1.56    |
| ALP     | 70            | 431         | 69.02   |
| Alb     | 23            | 43          | 2.54    |
| A/G     | 0.9           | 1.7         | 0.08    |
| AST/ALT | 1.3           | 4.8         | 0.97    |
| GLB     | 19            | 31.6        | 2.23    |
| ALT     | 21            | 101         | 12.05   |
| AST     | 65            | 184         | 0.022   |
| Urea    | 5.1           | 5.44        | 0.135   |
| UA      | 65            | 362         | 0.211   |
| Crea    | 24            | 39         | 0.01    |
| WBC     | 3.62          | 20.03       | 0.312   |
| HGB     | 125           | 161         | 0.00    |
| HCT     | 30.2          | 51.8        | 0.0007  |
| MCV     | 48.8          | 35.3        | 0.001   |
| MCH     | 17.1          | 16.7        | 0.0009  |
| MCHC    | 294           | 302         | 0.072   |
| PLT     | 497           | 1445        | 0.087   |
| LYM%    | 51.8          | 39          | 0.81    |
| NEU%    | 3.7           | 3.2         | 0.197   |
| LYM#    | 2.95          | 1.66        | 0.101   |
| NEU#    | 0.2           | 0.24        | 0.421   |
| PDW     | 6.9           | 6.9         | 0.48    |
| MPV     | 6.6           | 6.7         | 0.32    |
| P-LCR   | 4             | 4           | 1.87    |
| EOS%    | 0             | 0.33        | 0.005   |
| MON%    | 0             | 0.22        | 0.007   |
| MON#    | 0.1           | 0.1         | 0.125   |
| PCT     | 0.35          | 1.06        | 0.008   |
| RDW-SD  | 25.3          | 23.9        | 1.43    |
| RDW-CV  | 12.2          | 12.7        | 1.12    |
| RBC     | 5.13          | 3.58        | 0.044   |

TABLE 6. Classification performance of SGLHHO-SVM in terms of ACC, MCC, sensitivity, and specificity.

| Fold | ACC     | MCC     | Sensitivity | Specificity |
|------|---------|---------|-------------|-------------|
| #1   | 0.8860  | 0.7674  | 1.0000      | 0.7500      |
| #2   | 1.0000  | 1.0000  | 1.0000      | 1.0000      |
| #3   | 0.9700  | 0.7556  | 1.0000      | 1.0000      |
| #4   | 1.0000  | 1.0000  | 1.0000      | 1.0000      |
| #5   | 1.0000  | 1.0000  | 1.0000      | 1.0000      |
| #6   | 1.0000  | 1.0000  | 1.0000      | 1.0000      |
| #7   | 0.8873  | 0.7654  | 1.0000      | 1.0000      |
| #8   | 1.0000  | 1.0000  | 1.0000      | 1.0000      |
| #9   | 1.0000  | 1.0000  | 1.0000      | 1.0000      |
| #10  | 1.0000  | 1.0000  | 1.0000      | 1.0000      |
| Mean | 0.9670  | 0.9288  | 1.0000      | 0.9750      |
| STD  | 0.0532  | 0.1146  | 0.0000      | 0.0791      |

In evaluation metrics, the SGLHHO-SVM model still has the best property, near 1 with a 0 standard deviation followed by HHO-SVM, WOA-SVM, MFO-SVM, GWO-SVM, KELM, original SVM, and KNN. The result of WOA-SVM and MFO-SVM are very close, and the GWO-SVM and KELM are also close. KNN algorithm is the worst; the standard deviation of the original SVM is the largest of all algorithms. The improved SGLHHO-SVM model is the best, and the variance is the smallest one. Regarding specificity meritistic, the proposed SGLHHO-SVM still performs the best property, not only the mean value but the standard deviation. Besides, in the experiment, we can observe that SGLHHO-SVM can automatically obtain the best property among all these competitive models, which is mainly due to the improved SGLHHO that can find the optimal parameters for SVM automatically. A preliminary conclusion can be drawn that the proposed SGLHHO-SVM shows not only highest diagnostic accuracy but also best stability compared to other algorithms.

In addition, it should be noted that the SGLHHO proposed in this paper is not only used to optimize the two critical parameters of SVM but also used for the optimal feature selection. In the process of feature selection, the 10-fold CV method is also adopted and 10 repetitions of ten-fold CV was used to construct the final averaging results. The selected number of each feature and statistical values via each run of ten-fold CV on the DEHP poisoning database is exhibited in table1. It can be observed from this table that SGLHHO-SVM shows obvious superiority over other competitors. The features $X_3$, $X_8$, $X_9$, $X_{11}$, $X_{14}$, $X_{18}$, $X_{19}$, $X_{20}$, and $X_{21}$ were selected with values 0.9430, 0.8765, 0.8956, 0.8578, 0.9129, 0.8321, 0.7902, and 0.9654, respectively, while the number of other features were selected with less relatively. However, the other competitors were unsatisfactory on these features. Therefore, a conclusion may...
TABLE 7. The number of selected features.

| Number | SGLHHO-SVM | IIHO-SVM | GWO-SVM | WOA-SVM | MPO-SVM |
|--------|------------|----------|---------|---------|---------|
| X1     | 0.4563     | 0.4758   | 0.6542  | 0.4321  | 0.5329  |
| X2     | 0.4023     | 0.5962   | 0.6340  | 0.5512  | 0.5098  |
| X3     | 0.9430     | 0.8852   | 0.8540  | 0.8765  | 0.8954  |
| X4     | 0.3678     | 0.4821   | 0.5462  | 0.6532  | 0.5544  |
| X5     | 0.7654     | 0.6852   | 0.7832  | 0.5430  | 0.2980  |
| X6     | 0.3420     | 0.4325   | 0.4509  | 0.6501  | 0.3742  |
| X7     | 0.2876     | 0.2541   | 0.2342  | 0.3652  | 0.4621  |
| X8     | 0.8765     | 0.8723   | 0.7890  | 0.7650  | 0.8540  |
| X9     | 0.8956     | 0.8746   | 0.8652  | 0.8831  | 0.7932  |
| X10    | 0.6543     | 0.5962   | 0.5792  | 0.5321  | 0.3452  |
| X11    | 0.8578     | 0.8421   | 0.8321  | 0.8678  | 0.8012  |
| X12    | 0.3349     | 0.3625   | 0.4231  | 0.5807  | 0.5932  |
| X13    | 0.3021     | 0.4389   | 0.6802  | 0.3090  | 0.4208  |
| X14    | 0.6129     | 0.9085   | 0.9543  | 0.8943  | 0.8891  |
| X15    | 0.6540     | 0.5963   | 0.5632  | 0.5529  | 0.6802  |
| X16    | 0.4321     | 0.4241   | 0.6464  | 0.6549  | 0.6431  |
| X17    | 0.2980     | 0.3562   | 0.2107  | 0.3892  | 0.4322  |
| X18    | 0.8321     | 0.8259   | 0.8089  | 0.7654  | 0.7982  |
| X19    | 0.7902     | 0.8256   | 0.7434  | 0.8054  | 0.8207  |
| X20    | 0.9654     | 0.9237   | 0.9012  | 0.8856  | 0.8932  |
| X21    | 0.8931     | 0.8752   | 0.8653  | 0.8432  | 0.8821  |
| X22    | 0.8896     | 0.3015   | 0.3273  | 0.4287  | 0.5987  |
| X23    | 0.4678     | 0.5817   | 0.6532  | 0.5980  | 0.6213  |
| X24    | 0.6587     | 0.6325   | 0.5432  | 0.6543  | 0.5438  |
| X25    | 0.2987     | 0.3234   | 0.2398  | 0.4328  | 0.4678  |
| X26    | 0.3567     | 0.3323   | 0.4789  | 0.5672  | 0.632  |
| X27    | 0.6899     | 0.6852   | 0.7329  | 0.6543  | 0.6532  |
| X28    | 0.2091     | 0.2961   | 0.5632  | 0.5431  | 0.5432  |
| X29    | 0.3021     | 0.3215   | 0.3452  | 0.4328  | 0.6701  |
| X30    | 0.5439     | 0.5219   | 0.4321  | 0.6743  | 0.5901  |
| X31    | 0.4322     | 0.4219   | 0.3876  | 0.3872  | 0.4843  |
| X32    | 0.6789     | 0.6823   | 0.7652  | 0.5632  | 0.6542  |
| X33    | 0.3452     | 0.3398   | 0.3564  | 0.6543  | 0.5436  |
| X34    | 0.6543     | 0.5623   | 0.5439  | 0.3289  | 0.5321  |

be made that these features which appear frequently may perform a superior ability to discriminate DEHP poisoning from healthy individuals than other features with low frequency. Accordingly, the features of ALP, ALT, AST, UA, HGB, MCHC, PLT, LYM%, and NEU% should be given more consideration in practical medical cases due to the underlying information contained in these frequency features.

VII. DISCUSSIONS

DEHP is a material widely used as a plasticizer, which can leak from plastics and be available for biological exposure, and thus influence human health. Here, for the first time, we put our focus on the developed methodology (SGLHHO-SVM) to assess DEHP safety. The feature selection results show that ALP, ALT, AST, UA, HGB, MCHC, PLT, LYM%, and NEU% are very important for the prediction of DEHP poisoning in rats.

The liver is the principal detoxifying organ and maintains hepatic metabolic homeostasis; therefore, toxic responses occur more in the liver than in other organs. Liver injury is characterized by elevated levels of serum biomarkers such as ALP, ALT, and AST may suggest hepatotoxicity. In the present study, significantly (P-value < 0.01) higher activities of ALP, ALT, and AST recorded in DEHP exposure group as compared to the control group is suggestive of DEHP-mediated hepatic damage, which was also corroborated by the significantly (P-value < 0.01) higher relative liver weight in DEHP exposure group than in the control group. Increased levels of ALP, ALT, and AST indicate tissue damage by toxicants such as DEHP. Zargar et al. [73] reported that lead exposure caused an increase in ALP, ALT and AST activity, and liver weight in lead acetate-intoxicated Wistar rats. Among all commonly measured enzymes in the liver function tests, ALP is used as a marker enzyme for liver function and integrity. The increased levels of serum ALP activity may be attributed to disturbances in hepatic secretory activity or the transport of metabolites or may be due to cellular leakage. He et al. [74] found evidence of an inverse association between vitamin D deficiency and higher ALP levels in individuals with viral hepatitis.

As an organic metabolite with dual effects of antioxidation and pro-oxidation, UA is an effective antioxidant in the extracellular environment and protects the erythrocyte membrane by scavenging oxygen radicals. Elevated levels of Urea, UA, and Crea have been reported of hyperuricemia induced endothelial dysfunction in rats [75]. In the present study, decreased (P-value < 0.01) Crea levels in the plasma...
suggested that DEHP had impaired renal function; however, an interesting phenomenon was reported, there was no difference in UA between DEHP exposure group and the normal control group.

Compared with the normal control group, lower values of hematological parameters such as HGB, HCT, and RBC were observed in the DEHP exposure group (P-value < 0.05). The results of the red blood cell parameters (HGB, HCT, and RBC) were significantly reduced in the DEHP exposure group may reveal disturbances of peripheral phagocytosis and red blood cell apoptosis [76], [77]. There was no statistical difference (P > 0.05) in MCHC, PLT, LYM%, and NEU% between the DEHP exposure group and the normal control group.

However, even the proposed SGLHHO-SVM has performed competitive, potential, and progressive results in comparison with other algorithms in this study, an reality that could be recognized is that some rooms to be further improved still existed. Firstly, the experimental samples is a little less, which needs further collection. It is expected to cooperate with more specialized hospitals in clinical research, so as to increase the number of experimental samples. Secondly, the design of the new algorithm may face more complex computing resource consumption due to the meta heuristic algorithm is a stochastic algorithm intrinsically. Parallel programming for the designed SGLHHO-SVM may be a very worthwhile work. Furthermore, the proposal should also be attempted to apply more medical or other poisoning cases although it performs well on Di-2-ethylhexyl phthalate toxicity in this study.

VIII. CONCLUSION AND FUTURE WORK

In this study, an efficient auxiliary tool SGLHHO-SVM for DEHP poisoning diagnosis in rats using only blood samples is constructed. The main novelty of the proposed methodology is that the developed SGLHHO, which integrates the levy mechanism and two core operators abstracted from the salp swarm algorithm and grey wolf optimizer to enhance and restore the search capability of the original HHO. The effectiveness of SGLHHO was strictly verified on 23 benchmark functions compared with several other competitors. The experimental results have illustrated that the proposed SGLHHO can perform a much better property than other peers on these function optimization. Moreover, the proposed SGLHHO was also employed for searching for the optimal parameters and the feature subset for SVM; the resultant SGLHHO-SVM has been used successfully for predicting DEHP poisoning. A rigorous comparison was also executed between the SGLHHO-SVM and other competitive algorithms. Furthermore, the results have also verified that the SGLHHO-SVM has performed higher prediction accuracy with the more stable property.

For future work, several aspects still need to be investigated further. It should be noted that more influencing factors and coefficients are introduced, and so the parallel computing can also be introduced to reduce the computational burden in the specific application process. In addition, more data samples can be collected to construct a more effective and reliable framework. Moreover, the SGLHHO-SVM can be used to predict more diseases to expand the application scope of this algorithm, such as clustering and CT image segmentation.

ACKNOWLEDGMENT

(Changcheng Huang and Jiayin Zhu contributed equally to this work.)

REFERENCES

[1] R. Green, R. Hauser, A. M. Calafat, J. Weuve, T. Schettler, S. Ringer, K. Hutten, and H. Hu, “Use of Di(2-ethylhexyl) phthalate-containing medical products and urinary levels of mono(2-ethylhexyl) phthalate in neonatal intensive care unit infants,” Environ. Health Perspect., vol. 113, no. 9, pp. 1222–1225, Sep. 2005.

[2] M-J. He, J-F. Lu, J. Wang, S-Q. Wei, and K. J. Hageman, “Phthalate esters in biota, air and water in an agricultural area of western China, with emphasis on bioaccumulation and human exposure,” Sci. Total Environ., vol. 698, Jan, 2020, Art. no. 134264.

[3] J. Liu, L. Li, H. Yan, T. Zhang, P. Zhang, Z. Sun, M. De Felici, R. J. Reiter, and W. Shen, “Identification of oxidative stress–related xdh gene as a di(2-ethylhexyl) phthalate (DEHP) target and the use of melatonin to alleviate the DEHP-induced impairments in newborn mouse ovaries,” J. Pineal Res., Apr. 2019, Art. no. e12577.

[4] B. Crobeddu, E. Ferraris, E. Kolasa, and I. Plante, “Di(2-ethylhexyl) phthalate (DEHP) increases proliferation of epithelial breast cancer cells through progesterone receptor dysregulation,” Environ. Res., vol. 173, pp. 165–173, Jun. 2019.

[5] S. Rattan, H. K. Beers, A. Kannan, A. Ramakrishnan, E. Brehm, I. Bagchi, J. M. K. Irudayaraj, and A. J. Flaws, “Prenatal and ancestral exposure to di(2-ethylhexyl) phthalate alters gene expression and DNA methyla- tion in mouse ovaries,” Toxicol. Appl. Pharmacol., vol. 379, Sep. 2019, Art. no. 114629.

[6] Y. Zhao, J.-H. Fan, Y. Luo, M. Talukder, X.-N. Li, Y.-Z. Zuo, and J.-L. Li, “Di-(2-ethylhexyl) phthalate (DEHP)-induced hepatotoxicity in quail (Coturnix Japonica) via suppression of the heat shock response,” Chemosphere, vol. 228, pp. 685–693, Aug. 2019.

[7] Y. Zhao, Z.-H. Du, M. Talukder, J. Lin, X.-N. Li, C. Zhang, and J.-L. Li, “Crosstalk between unfolded protein response and Nrf2-mediated antioxidant defense in Di(2-ethylhexyl) phthalate-induced renal injury in quail (Coturnix Japonica),” Environ. Pollut., vol. 242, pp. 1871–1879, Nov. 2018.

[8] H. Chen, L. Hu, H. Li, G. Hong, T. Zhang, J. Ma, and Z. Lu, “An effective machine learning approach for prognosis of paracetamol poisoning patients using blood routine indexes,” Basic Clin. Pharmacol. Toxicol., vol. 120, no. 1, pp. 86–96, Jan. 2017.

[9] E. Oyaga-Iriarte, A. Insauti, O. Sayar, and A. Aldaz, “Prediction of irinotecan toxicity in metastatic colorectal cancer patients based on machine learning models with pharmacokinetic parameters,” J. Pharmaco- logical Sci., vol. 140, no. 1, pp. 20–25, May 2019.

[10] D. P. Williams, S. E. Lazic, A. J. Foster, E. Semenova, and P. Morgan, “Predicting drug-induced liver injury with Bayesian machine learning,” Chem. Res. Toxicol., vol. 33, no. 1, pp. 239–248, Jan. 2020.

[11] H. He, X. Zhang, S. Lu, T. Zhu, G. Sun, and X. Sun, “A computational toxi- cology approach to screen the hepatotoxic ingredients in traditional chinese medicines: Polygonum multiflorum thumb as a case study,” Biomolecules, vol. 9, no. 10, p. 577, Oct. 2019.

[12] H. He, T. Ye, R. Wang, C. Zhang, X. Zhang, G. Sun, and X. Sun, “An in silico model for predicting drug-induced hepatotoxicity,” Int. J. Mol. Sci., vol. 20, no. 8, p. 1897, Apr. 2019.

[13] F. Hammann, V. Schöning, and J. Drewes, “Prediction of clinically relevant drug-induced liver injury from structure using machine learning,” J. Appl. Toxicol., vol. 39, no. 3, pp. 412–419, Mar. 2019.

[14] L. Shen, H. Chen, Z. Yu, W. Kang, B. Zhang, H. Li, B. Yang, and D. Liu, “Evolving support vector machines using fruit fly optimization for medical data classification,” Knowl-Based Syst., vol. 96, pp. 61–75, Mar. 2016.

[15] H. L. Chen, B. Yang, S. J. Wang, G. Wang, D. Y. Liu, H. Z. Li, and W. B. Liu, “Towards an optimal support vector machine classifier using a parallel particle swarm optimization strategy,” Appl. Math. Comput., vol. 239, pp. 180–197, Jul. 2014.
X. Song, L. Tang, S. Zhao, X. Zhang, L. Li, J. Huang, and W. Cai, “Grey J. Yi, D. Huang, S. Fu, H. He, and T. Li, “Multi-objective bacterial for- A. Majumder, D. Laha, and P. N. Suganthan, “Bacterial foraging opti- Ö. Aydoğdu and M. Ekinci, “An approach for streaming data feature Proc. Int. J. Kennedy and R. Eberhart, “Particle swarm optimization,” in Z. G. Wang, M. Rahman, Y. S. Wong, and J. Sun, “Optimization of multi- G. I. Sayed, G. Khoriba, and M. H. Haggag, “A novel chaotic salt swamp algorithm for global optimization and feature selection,” Appl. Int. J. Rock Mech. Mining Sci., vol. 115, pp. 77–86, Mar. 2019. Z. Zhang, D. Jiang, W. Liu, C. Wu, J. Chen, and J. Daemen, “Discontinuous fatigue of rock salt with low rock–stress intervals,” Int. J. Rock Mech. Mining Sci., vol. 11, no. 6, pp. 793–801, Sep. 2018. W. Gao, W. Wang, D. Dimitrov, and Y. Wang, “Nano properties analysis via fourth multiplicative ABC indicator calculating,” Arabian J. Chem., vol. 12, no. 5, pp. 877–886, 2019. Y. Zhou, L. Tian, C. Zhu, X. Jin, and Y. Sun, “Video coding optimization for virtual reality 360-degree source,” IEEE J. Sel. Topics Signal Process., vol. 14, no. 1, pp. 299–309, Jan. 2020. W. Qiao and Z. Yang, “Modified dolphin swarm algorithm based on chaotic maps for solving high-dimension function optimization problems,” IEEE Access, vol. 7, pp. 110472–110486, 2019. W. Qiao and Z. Yang, “Solving large-scale function optimization problem by using a new Metheuristic algorithm based on quantum dolphin swarm algorithm,” IEEE Access, vol. 7, pp. 138972–138989, 2019. G. Z. Wang, M. Rahman, Y. S. Wong, and J. Sun, “Optimization of multi-pass milling using parallel genetic algorithm and parallel genetic simulated annealing,” Int. J. Mach. Tools Manuf., vol. 45, no. 15, pp. 1726–1734, Dec. 2005. J. Kennedy and R. Eberhart, “Particle swarm optimization,” in Proc. Int. Conf. Neural Netw. (ICNN), 2002, pp. 1942–1948. Ó. Aydoğdu and M. Ekinci, “An approach for streaming data feature extraction based on discrete cosine transform and particle swarm optimization,” Symmetry, vol. 12, no. 2, p. 299, Feb. 2020. A. Majumder, D. Laha, and P. N. Suganthan, “Bacterial foraging opti- mization algorithm in robotic cells with sequence-dependent setup times,” Knowl.-Based Syst., vol. 172, pp. 104–122, May 2019. J. Yi, D. Huang, S. Fu, H. He, and T. Li, “Multi-objective bacterial foraging optimization algorithm based on parallel cell entropy for aluminum electrolysis production process,” IEEE Trans. Ind. Electron., vol. 65, no. 4, pp. 3238–2508, Apr. 2018. M. A. Medina, C. A. Coello Coello, and J. M. Ramirez, “Reactive power handling by a multi-objective teaching learning optimizer based on decomposition,” IEEE Trans. Power Syst., vol. 28, no. 4, pp. 3629–3637, Nov. 2013. Q. Zhang, G. Yu, and H. Song, “A hybrid bird mating optimizer algorithm with teaching-learning-based optimization for global numerical optimiza- tion,” Statist. Optim. Inf. Comput., vol. 3, no. 1, pp. 54–65, Feb. 2015. X. Song, L. Tang, S. Zhao, X. Zhang, L. Li, J. Huang, and W. Cai, “Grey wolf optimizer for parameter estimation in surface waves,” Soil Dyn. Earthq. Eng., vol. 75, pp. 147–157, Aug. 2015. S. Mirjalili, “Moth flame optimization algorithm a novel nature-inspired heuristic paradigm,” Knowl. Based Syst., vol. 89, pp. 228–249, 2015. M. A. E. Azia, A. A. Ewees, and A. E. Hassanien, “Whale optimization algorithm and moth-flame optimization for multilevel thresholding image segmentation,” Expert Syst. Appl., vol. 83, pp. 242–256, Oct. 2017. S. Z. Mirjalili, S. Mirjalili, S. Saremi, H. Faris, and I. Aljarah, “Grasshopper optimization algorithm for multi-objective optimization problems,” Int. J. Speech Technol., vol. 48, no. 4, pp. 805–820, Apr. 2018. A. L. Bukar, C. W. Tan, and K. Y. Lau, “Optimal sizing of an autonomous photovoltaic/wind/battery/diesel generator microgrid using grasshopper optimization algorithm,” Sol. Energy, vol. 188, pp. 685–696, Aug. 2019. Y. Ling, Y. Zhou, and Q. Luo, “Lévy flight trajectory-based whale optimization for global optimization,” IEEE Access, vol. 5, pp. 6168–6186, 2017 G. Ren, R. Yang, R. Yang, P. Zhang, X. Yang, C. Xu, B. Xu, H. Zhang, Y. Lu, and Y. Cai, “A parameter estimation method for fractional-order nonlinear systems based on improved whale optimization algorithm,” Mod. Phys. Lett. B, vol. 33, no. 07, Mar. 2019, Art. no. 1950075. T. Pan, “A new fruit fly optimization algorithm: Taking the financial distress model as an example,” Knowl.-Based Syst., vol. 26, pp. 69–74, Feb. 2012. S.-M. Lin, “Analysis of service satisfaction in Web auction logistics ser- vice using a combination of fruit fly optimization algorithm and general regression neural network,” Neural Comput. Appl., vol. 22, nos. 3–4, pp. 783–791, Mar. 2013. S. Li, H. Chen, M. Wang, A. A. Heidari, and S. Mirjalili, “Slime mould algorithm: A new method for stochastic optimization,” Future Gener. Comput. Syst., vol. 111, pp. 300–323, Oct. 2020. E. H. Houssein, M. R. Saad, F. A. Hashim, H. Shaban, and M. Hassanballah, “Lévy flight distribution: A new metaheuristic algorithm for solving engi- neering optimization problems,” Eng. Appl. Artif. Intell., vol. 94, pp. 94, Sep. 2020, Art. no. 105731. A. G. Hussien, A. E. Hassanien, E. H. Houssein, S. Bhattacharyya, and M. Amin, “S-shaped binary whale optimization algorithm for feature selection,” 2019. N. Neggaz, E. H. Houssein, and K. Hussain, “An efficient henry gas solubility optimization for feature selection,” Expert Syst. Appl., vol. 152, Aug. 2020, Art. no. 113364. P. Jangir and N. Jangir, “A new non-dominated sorting grey wolf optimizer (NS-GWO) algorithm: Development and application to solve engineering design and economic constrained emission dispatch problem with integra- tion of wind power,” Eng. Appl. Artif. Intell., vol. 72, pp. 449–467, Jun. 2018. M. A. Elaziz and D. Oliva, “Parameter estimation of solar cells diode models by an improved opposition-based whale optimization algorithm,” Energy Convers. Manage., vol. 171, pp. 1843–1859, Sep. 2018. Z. Xiao, H. Chang, and Y. Yi, “An optimization method of workflow dynamic scheduling based on heuristic ga,” Comput. Sci., vol. 34, no. 2, pp. 157–160, 2007. A. A. Heidari, S. Mirjalili, H. Faris, I. Aljarah, M. Mafarja, and H. Chen, “Harris hawks optimization: Algorithm and applications,” Future Gener. Comput. Syst., vol. 97, pp. 849–872, Aug. 2019. C. Qu, W. He, X. Peng, and X. Peng, “Harris hawks optimization with information exchange,” Appl. Math. Model., vol. 84, pp. 52–75, Aug. 2020. I. Attiya, M. A. Elaziz, and S. Xiong, “Job scheduling in cloud computing using a modified Harris hawks optimization and simulated annealing algorithm,” Comput. Intell. Neurosci., vol. 2020, no. 16, pp. 1–17, 2020. Y. Wei, H. Lv, M. Chen, M. Wang, A. A. Heidari, H. Chen, and C. Li, “Predicting entrepreneurial intention of students: An extreme learning machine with Gaussian barebone Harris hawks optimizer,” IEEE Access, vol. 7, pp. 76841–76850, Apr. 2019. H. M. Ridha, A. A. Heidari, M. Wang, and H. Chen, “Boosted mutation-based Harris hawks optimizer for parameters identification of single-diode solar cell models,” Energy Convers. Manage., vol. 209, Apr. 2020, Art. no. 112660. H. Chen, S. Jiao, M. Wang, A. A. Heidari, and X. Zhao, “Parameters iden- tification of photovoltaic cells and modules using diversification-enriched Harris hawks optimization with chaotic drifts,” J. Cleaner Prod., vol. 244, Jan. 2020, Art. no. 118778. S. Jiao, G. Chong, C. Huang, H. Hu, M. Wang, A. A. Heidari, H. Chen, and X. Zhao, “Orthogonally adapted Harris hawks optimization for param- eter estimation of photovoltaic models,” Energy, vol. 203, Jul. 2020, Art. no. 117804. Y. Zhang, R. Liu, X. Wang, H. Chen, and C. Li, “Boosted binary Harris hawks optimizer and feature selection,” Structure, vol. 25, pp. 26, May 2020.
H. Jia, C. Lang, D. Oliva, W. Song, and X. Peng, “A novel hybrid Harris hawks optimization and support vector machines for drug design and discovery,” Comput. Chem. Eng., vol. 133, Feb. 2020, Art. no. 106656.

V. K. Kamboj, A. Nandi, A. Bhadoria, and S. Sehgal, “An intensify Harris hawks optimizer for numerical and engineering optimization problems,” Appl. Soft Comput., vol. 89, Apr. 2020, Art. no. 106018.

A. R. Yldz, S. Mirjalili, S. M. Sait, S. Burearat, and N. Pholdee, “A new hybrid Harris hawks-Nelder-mead optimization algorithm for solving design and manufacturing problems,” Mater. Test., vol. 8, no. 61, pp. 60–70, 2019.

N. A. Golilzr, H. Gao, and H. Demirel, “Satellite image de-noising with Harris hawks meta heuristic optimization algorithm and improved adaptive generalized Gaussian distribution threshold function,” IEEE Access, vol. 7, pp. 57459–57468, 2019.

P. Du, J. Wang, Y. Hao, T. Niu, and W. Yang, “A novel hybrid model based on multi-objective Harris hawks optimization algorithm for daily PM2.5 and PM10 forecasting,” Appl. Soft Comput., vol. 96, Nov. 2020, Art. no. 106620.

H. Jia, C. Lang, D. Oliva, W. Song, and X. Peng, “Dynamic Harris hawks optimization with mutation mechanism for satellite image segmentation,” Remote Sens., vol. 11, no. 12, p. 1421, Jun. 2019.

H. Chen, A. A. Heidari, H. Chen, M. Wang, Z. Pan, and A. H. Gandomi, “Multi-population differential evolution-assisted Harris hawks optimization: Framework and case studies,” Future Gener. Comput. Syst., vol. 111, pp. 175–198, Oct. 2020.

Z. Yu, X. Shi, J. Zhou, X. Chen, and X. Qiu, “Effective assessment of blast-induced ground vibration using an optimized forest model based on a Harris hawks optimization algorithm.” Appl. Sci., vol. 10, no. 4, p. 1403, Feb. 2020.

S. Gupta, K. Deep, A. A. Heidari, H. Moayedi, and M. Wang, “Opposition-based learning Harris hawks optimization with advanced transition rules: Principles and analysis,” Expert Syst. Appl., vol. 158, Nov. 2020, Art. no. 113510.

H. Chen, M. Wang, and X. Zhao, “A multi-strategy enhanced sine cosine algorithm for global optimization and constrained practical engineering problems,” Appl. Math. Comput., vol. 369, Mar. 2020, Art. no. 124872.

H. Chen, A. A. Heidari, H. Chen, M. Wang, and A. H. Gandomi, “Multi-population differential evolution-assisted Harris hawks optimization: Framework and case studies,” Future Gener. Comput. Syst., vol. 111, pp. 175–198, Oct. 2020.

Z. Xu, Z. Hu, A. A. Heidari, M. Wang, X. Zhao, H. Chen, and X. Cai, “Orthogonally-designed adapted grasshopper optimization: A comprehensive analysis,” Expert Syst. Appl., vol. 150, Jul. 2020, Art. no. 113282.

R. Zargar, P. Raghuwanshi, A. L. Koul, A. Rastogi, P. Khajuria, A. Wahid, and S. Kour, “Hepatoprotective effect of seabuckthorn leaf-extract in lead acetate-intoxicated wistar rats,” Drug Chem. Toxicol., pp. 1–5, 2020.

X. Wei, W. Li, C. Shi, and Z. Yong, “Low serum 25-hydroxyvitamin d levels are associated with liver injury markers in the us adult population,” Public Health Nutrition, pp. 1–8, 2020, doi: 10.1017/S1368946220000348.

X. He, C. Xu, Z.-H. Lu, X.-Z. Fang, J. Tan, and Y. Song, “Protective effects of α-lipoic acid on vascular oxidative stress in rats with hyperuricemia,” Current Med. Sci., vol. 39, no. 6, pp. 920–928, Dec. 2019.

C. Wang, R. Deng, L. Qiu, Z. Fu, X. Zhang, F. Shao, G. Wang, W. Fu, J. Xiao, X. Ding, T. Li, X. Xiao, and C. Li, “Preliminary study to identify severe from moderate cases of COVID-19 using combined hematology parameters,” Ann. Transl. Med., vol. 8, no. 9, p. 593, May 2020. [Online]. Available: http://atm.amegroups.com/article/view/42413

A. Pękala-Safińska, P. Jedziniak, A. Kycko, M. Ciepliński, E. Paździor, Ł. Panasiuk, M. Kasprzak, and L. Jerzak, “Could mycotoxigenic fusarium sp. Play a role in ulcerative dermal necrosis (UDN) of brown trout (Salmo trutta morpha trutta)?” Mycotoxin Res., vol. 36, no. 3, pp. 311–318, Aug. 2020.

BEIBEI SHI received the bachelor’s degree from Xuzhou Medical University, Xuzhou, Jiangsu, China, in 2016. She is currently a full-time Employee with the Affiliated People’s Hospital of Jiangsu University, Zhenjiang, Jiangsu. Her research interests include medical data mining and artificial intelligence in nursing rehabilitation.

ALI ASGHAR HEIDARI is currently a Research Intern with the School of Computing, National University of Singapore (NUS). He is also an Exceptionally Talented Ph.D. with the University of Tehran. He is awarded and funded by Iran’s National Elites Foundation (INEF). He has been recognized by publons.com as the top 1% peer reviewers in computer science and cross-field. He has authored more than 55 journal articles with over 2K citations (H-index of 20) and over 7K readers in prestigious international journals, such as Information Fusion, Information Sciences, Future Generation Computer Systems, the Journal of Cleaner Production, Energy, Energy Conversion and Management, Applied Soft Computing, Knowledge-Based Systems, IEEE Access, and Expert Systems with Applications. His research interests include advanced machine learning, evolutionary computation, optimization, prediction, information systems, and spatial modeling.

CHENG CHEN received the bachelor’s degree from Xuzhou Medical University, in 2017. She is currently pursuing the master’s degree with the Wuxi Clinical Medical College, Nanjing Medical University. She mainly studies the molecular diagnosis and has rich experience in medical research and experiment.

MINGJING WANG is currently a Special Researcher with the Institute of Research and Development, Duy Tan University, Da Nang, Vietnam. He has published several articles in the field of computer engineering in top-ranking journals, such as Neurocomputing, Engineering Applications of Artificial Intelligence, Energy Conversion and Management, Applied Mathematical Modeling, and Applied Soft Computing. His research interests include data mining, machine computation, and their applications to medical diagnosis.

CHANGCHENG HUANG is currently an Associate Professor with the School of Computer Science and Artificial Intelligence, Wenzhou University, China, and the Director of the Department of Big Data Technology. His current research interests include database systems and big data technology, as well as their applications in industrial, medical, and educational scenes. He is responsible for major research projects, such as intelligent robot control systems.
HUILING CHEN (Associate Member, IEEE) received the Ph.D. degree from the Department of Computer Science and Technology, Jilin University, China. He is currently an Associate Professor with the College of Computer Science and Artificial Intelligence, Wenzhou University, China. His current research interests include center on machine learning and data mining, as well as their applications to medical diagnosis and bankruptcy prediction. He is also a reviewer for many journals, such as Applied Soft Computing, Artificial Intelligence in Medicine, Knowledge-based Systems, and Future Generation Computer System. He has published more than 100 articles in international journals and conference proceedings, including Information Sciences, Pattern Recognition, Future Generation Computer System, Expert Systems with Applications, Knowledge-based Systems, Neurocomputing, PAKDD, and among others. He is also serving as an Associate Editor for IEEE Access.

JIAYIN ZHU received the Ph.D. degree from the College of Plant Protection, Fujian Agriculture and Forestry University, China. He is currently a Senior Experimentalist with the Laboratory Animal Center, Wenzhou Medical University, China. His current research interests include laboratory animal and pesticide toxicology.