Int. J. Nonlinear Anal. Appl. 13 (2022) 1, 3441-3450
ISSN: 2008-6822 (electronic)
http://dx.doi.org/10.22075/ijnaa.2022.6107

Hyperparameters optimization of support vector regression using black hole algorithm

Saifuldeen Dheyauldeen Alrefaee[a,*], Salih Muayad Al Bakal[a], Zakariya Yahya Algamal[b]

[a]Department of Operations Research and Intelligent Technologies, University of Mosul, Mosul, Iraq
[b]Department of Statistics and Informatics, University of Mosul, Mosul, Iraq

(Communicated by Madjid Eshaghi Gordji)

Abstract

The support vector regression (SVR) technique is considered the most promising and widespread way in the prediction process, and raising the predictive power of this technique and increasing its generalization ability well depends on tunning its hyperparameters. Nature-inspired algorithms are an important and effective tool in optimizing or tuning hyperparameters for SVR models. In this research, one of the algorithms inspired by nature, the black hole algorithm (BHA), by adapting this algorithm to optimize the hyperparameters of SVR, the experimental results, obtained from working on two data sets, showed, the proposed algorithm works better by finding a combination of hyperparameters as compared to the grid search (GS) algorithm, in terms of prediction and running time. In addition, the experimental results show the improvement of the prediction and computational time of the proposed algorithm. This demonstrates BHA's ability to find the best combination of hyperparameters.

Keywords: Support Vector Regression (SVR), Black Hole Algorithm (BHA), Hyperparameters

1. Introduction

The support vector machine (SVM) is considered one of the most important techniques that have received great interest among many researchers in recent years, due to its many advantages and its outstanding performance in classification and regression [12, 25, 28]. Where this technique was first used to solve the classification problem through the use of Vapnik's ε - insensitive loss function, and

*Corresponding author
Email addresses: saifldeen.alrefaee@uomosul.edu.iq (Saifuldeen Dheyauldeen Alrefaee), salih.mooaed@uomosul.edu.iq (Salih Muayad Al Bakal), zakariya.algamal@uomosul.edu.iq (Zakariya Yahya Algamal)

Received: November 2021 *Accepted:* January 2022

then this technique was developed to remedy the regression problem where it was named support vector regression (SVR) [3, 21, 22, 25].

The performance of any SVR model depends on the optimal selection of hyperparameters, as the process of optimizing the hyperparameters has received the attention of many scientists and researchers to reach the robustness in SVR models by raising the predictive power of these models [7, 15, 16, 17, 24].

In a study presented by [7] on the workable chosen of hyperparameters for SVM regression through a proposed methodology indirect selection of parameters from training data instead of the re-sampling methods used in SVM applications, in addition to his focus on the significance of Vapnik's ε - insensitive loss function to regression problems with unlimited samples and superior generalization performance of the SVR, for limited sample settings. For quantitative spectral analysis, [24] presented a new method for improving the hyperparameters of SVR, depending on the genetic algorithm and the simplex algorithm, the study concluded that SVR is less sensitive to spectral changes compared to the widely applied partial least squares (PLS) method.

A study was presented to predict tourism demand through the creation of an effective SVR model and its comparison with back-propagation neural networks (BPNN) and the autoregressive integrated moving average (ARIMA) model, and the improvement of the hyperparameters of the SVR has been proposed using the genetic algorithm (GA), and the results proved the superiority of the proposed GA-SVR model over BPNN and ARIMA models [5].

Illustrated [15] the great interest in SVR support vector regression in chemical measurements and the importance of the correct selection of three hyperparameters in SVR modeling, they presented a proposal for the rapid improvement of the hyperparameters based on the grid search (GS) method, cross-validation (CV) and theoretical techniques, taking into consideration the predictive power of SVR models, it demonstrated the high predictive power of SVR models and the small computational time of the proposed method. Also [8] developed a new hybrid smart artificial firefly colony algorithm (SAFCA) specifically designed to solve engineering prediction problems, and SAFCA was used to improve or adjust the hyperparameters in the SVR model, as the proposed model achieved better error rates than other predictive methods.

[18] presented improved the hyperparameters of the support vector regression using one of the nature-inspired algorithms, the artificial bee colony algorithm, where they proved the superiority of the proposed method over the ordinary support vector regression. Also [17] showed the great influence of SVR hyperparameters on the prediction of SVR performance as it proved that using the sine and cosine algorithm (SCA) can find optimal values for the hyperparameters SVR, and thus improve generalization performance on unknown data.

A study presented by [16] on important uses of support vector regression (SVR) in chemical sensor applications, illustrated the importance of the correct selection of hyperparameters in SVR performance and proposed the use of the generalized pattern search algorithm (GPS) as a faster alternative to identifying hyperparameters compared to more complex alternatives. [1] Use the manual search (MS) method and the genetic algorithm (GA) search method to improve the hyperparameters of the developed SVR regression model, and it proved the accuracy of the estimation and the generalization of the developed models in measuring the level of environmental pollution.

The use of meta-heuristics algorithms inspired by nature has expanded greatly in recent years, including population-based algorithms, in solving optimization problems including hyperparameter tuning [8, 17]. [2] presented a particle swarm optimization method to determine the tuning parameter in PSVM with L1-standard penalty, which will help to find the most important descriptors in the formation of the Quantitative Structure-Activity Classification (QSAR) with high performance, where concluded that the proposed method showed favorable performance at a low sample size and

a high number of descriptors compared to other methods. Also [20] presented a hybridization between the Firefly algorithm (FFA) and the Modified Adomian Decomposition Method (MADM) to estimate several parameters of the nonlinear Hirota-Satsuma coupled KdV system to choose the best parameters.
The black hole algorithm (BHA), a new population-based optimization algorithm proposed by [14], and its basic idea are inspired by the phenomenon of a black hole, which is a region of space that can absorb anything close to its line of gravity, including light. A BHA starts with an initial set of stars for the optimization problem and an objective function that is computed for it [4, 14, 19].
This algorithm is characterized by its speed and high accuracy in optimizing the hyperparameters, as the results showed the superiority of this algorithm over the grid search algorithm by evaluation criteria: The Root Mean Squared Error and time test in the optimization process.

The main direction of this work is to offer a new BHA-SVR model to optimize SVR hyperparameters using a black hole algorithm, which improves regression performance.
The remainder of the paper was divided as follows: Section (2) includes the work related to the regression of the supporting vector SVR and the effect of its hyperparameters on its performance. Section 3 describes the black hole algorithm. Section 4 introduces the proposed BHA-SVR method. Section (5), includes the application side and results in addition to the data set in section (6). The final section explains the conclusions.

2. Support Vector Regression and hyperparameters Selection

2.1. Support Vector Regression

Support Vector Regression (SVR) is a machine learning technique presented by [25], the main idea is to construct a predictive linear model based on minimizing the complexity of the model and at the same time reducing experimental risks. In regression formulation, the machine learning methodology it's by introducing the ϵ-insensitive loss function, and this process is used to estimate an unknown continuous function based on a certain number of training samples set [18, 17].

Assume that there is a training data set as $D = \{(x_i, y_i), i = 1 : n\}$ where $x_i \in \mathbb{R}^d$ is the input vector which has d dimension, $y_i \in \mathbb{R}$ is the output vector and the size of a training data is denoted by N . SVR is being fit linear model by a function $f(x, w)$, where the input x is first mapped of the training data upon high-dimensional feature space using some fixed (nonlinear) mapping, and then a linear model is constructed in this feature space [6, 16].

$$f(x, w) = w.\varphi(x) + b \tag{2.1}$$

where $\varphi(x)$ denotes a mapping of x in high dimensional feature space, b is a bias vector, and w is a weight vector. Through a loss function, the accuracy of the estimate is measured to calculate the predictive error, where the support vector regression uses a new loss function called the ε -insensitive loss function, as defined below [7, 26].

$$L(y, f(x, w)) = \begin{cases} 0 & \text{if} \quad |y_i - f(x, w)| \leq \epsilon \\ |y_i - f(x, w)| - \epsilon & \text{Otherwise} \end{cases} \tag{2.2}$$

Note that when $\varepsilon = 0$, the ε -insensitive loss function coincides with the least-modulus loss function and with a special case of Huber's robust loss function, thus it will be important to compare the performance of the prediction of SVM with the regression estimates obtained by the least-modulus loss function [9].

SVR executes linear regression in the high-dimension using ε -insensitive loss function and also attempts to decrease model intricacy by minimizing w^2. To do that, SVR used the positive slack variables ξ_i ,ξ_i^* $i = 1, \ldots, n$ to determine the deviation of training data outer ε - insensitive tube, To express the convex optimization problem and make it as a function of minimization, so [6]

$$\begin{aligned} &\text{Minimize} \quad \frac{1}{2}w^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*) \\ &\text{Subject to} \quad \begin{cases} y_i - f(x,w) \le \varepsilon + \xi_i^* \\ f(x,w) - y_i \le \varepsilon + \xi_i \\ \xi_i \ , \xi_i^* \ge 0 \qquad i = 1, \ldots, n \end{cases} \end{aligned} \tag{2.3}$$

where the constant C > 0 known as a regularization hyperparameter which calculates the tradeoff among the model intricacy (flatness) of $f(x, w)$ and empirical risk. The optimization problem shown in Eq. (2.3) can be converted to a dual problem by Lagrange multipliers, also can apply Implementation the Karush–Kuhn–Tucker (KKT) restrictions.
To overcome this problem, the function will be [16]

$$\begin{aligned} \textit{Minimize} \quad & -\frac{1}{2}\sum_{i,j=1}^{n}(\alpha_i - \alpha_i^*)\left(\alpha_j - \alpha_j^*\right)\mathbf{k}(x_i, x) \\ & - \ \varepsilon\sum_{i=1}^{n}(\alpha_i - \alpha_i^*) + \ \sum_{i=1}^{n} y_i(\alpha_i - \alpha_i^*) \\ \textit{Subject to} \quad & \sum_{i=1}^{n}(\alpha_i - \alpha_i^*) = 0 \quad \textit{and} \quad 0 \le \alpha_i, \ \alpha_i^* \le C \end{aligned} \tag{2.4}$$

By using kernel trick, the optimization problem can be reformulated based on the solution above as a regression function [27]

$$f(x) = \sum_{i=1}^{n}(\alpha_i - \alpha_i^*)\mathbf{k}(x_i, x) + b \tag{2.5}$$

Where $\alpha_i, \ \alpha_i^*$ are Lagrange multipliers, and $\mathbf{k}(x_i, x)$ refers to the kernel function.

2.2. Hyperparameters Selection for SVR

There are many methods for optimal selection of hyperparameters, and the default of these methods is the grid search, there are many nature-inspired algorithms used for tuning hyperparameters, the most we can talk about: Particle Swarm [2],[18] and grasshopper algorithm [3]. In this paper, we will use the BHA for choosing the appropriate hyperparameters of SVR, where this method has many features, which we will learn more about it in Section (3). First of all, we must know how each of these hyperparameters is affected.

It is known that the high accuracy of the estimation and the good performance of SVR depends mainly on the good selection of hyperparameters which consists of both C, ε and the kernel function parameter σ. As for the selection of the hyperparameter C, this leads to an increase in the number of support vectors which in turn increases the estimation efficiency for all training data, while selecting a lower value for C tend to the model flattest.

The hyperparameter ε it's used to fit the training data and control the width of the ε -insensitive tube, as the value of this parameter can directly affect the number of support vectors used in the

regression function. If the value chosen for ε is large, this leads to the determination of a small number of support vectors, which increases the flatness of the model, while choosing a smaller value for this parameter increases the number of support vectors that effect the complexity of the model.

As for the value of the hyperparameter σ, if it is small or large, this leads to a large effect of the support vectors or an increase in the complexity of the model (increased flattening), so there is a different effect for each hyperparameter on the model complexity of the feature space. Therefore, we can use new metaheuristic algorithms to find the optimum combination of hyperparameters that ensure a good model [7, 16].

3. Black Hole Algorithm (BHA)

The BHA is the newest sophisticated optimization algorithm and it's a method inspired that depicts the demeanor of BH in space by a mathematical expression and terms [19], where considered it's one of the strangest and most fascinating things found in outer space. The BHs are objects of high density that have a great force of gravity that even light cannot escape from their fist if it gets close enough, The region called the event horizon which stars are not allowed to escape from it [23]. The BHA is distinctive and popular which considered as one of the nature-inspired algorithms. This algorithm works to develop the population that has been created in the form of a group of stars to solve a specific problem and distributed them randomly in the sample space of the search space to reach the optimal solution by definite techniques [14, 19]. In the BHA a developing of the population is done by moving all the stars towards the better star in each iteration, namely a BH, and replacing these stars (suggest a solution) that enter within the range of the BH by newly generated stars (suggest a solution) in the search space [10].

After the initializing of BH and stars, evaluate the fitness values of the population. The selected star who is the representative of the better fitness value will be the center and can ingest the others. After the initialization process, the black hole begins the process of absorbing stars around it and all-stars begin to move towards it, and the process of star absorption by the BH will be formulated as:

$$x_i\,(t+1) = x_i\,(t) + rand \times (x_{\mathrm{BH}} - x_i\,(t)) \quad ; i = 1, 2, 3, \ldots, N \tag{3.1}$$

where $x_i\,(t)$ and $x_i\,(t+1)$ are the locations of the ith star at iterations t and $t+1$, respectively. x_{BH} is the location of a BH in the search space. rand is a random number in the interval [0,1]. N is the number of stars. While moving the stars towards a BH, a star may reach a location with the highest fitness than the BH, or vice versa the BH moves to a location of that star. The BHA will continue with the new location of a BH and then stars are starting to move to this new location [4, 13]. Furthermore, there's a chance of passing the event horizon during the move of stars towards the BH. Every star (suggest a solution) moving very close or passes the event horizon of the BH will be ingested by a BH, and is disappeared everlastingly. As in this case, a new star (suggest a solution) is generated and distributed randomly in the search space and starts a new search [19]. That is achieved to keep the number of stars constant [4]. After all the stars have been moved, then the next iteration will take place. The radius (R) of the event horizon in the BHA is calculated as:

$$R = \frac{f_{\mathrm{BH}}}{\sum_{i=1}^{N} f_i} \tag{3.2}$$

where f_{BH} is the fitness value of the BH, f_i is the fitness value of the ith star and N the number of stars. If the distance between any star and the BH is less than R, then the star is absorbed. In such a situation, a new star is created and distributed randomly in the search space [11, 14]. Based on

the previous description the choosing of random locations for the BHA is the first step to Initialize a population of stars in the search space, this algorithm will be illustrated as follows:

With a random location, Initialize a population of stars in the search space

Loop

1. The first step is to calculate the objective function for each star
2. Select the best star as a BH if it has the best fitness value
3. Change the position of the stars according to Equation (3.1)
4. exchange the position of a BH with any star arrive a position with lower cost from the BH
5. If a star passed the event horizon of the black hole, change it with a new star in a random position in a space
6. If the suspension norm is met (high values of fitness satisfactorily), exit the loop

End loop

Through the foregoing, two important points of the black hole algorithm can be identified, namely, the simplicity of implementation and the fact that it is free from the issues of hyperparameters tuning [4].

4. BHA in determining SVR hyperparameters (BHA-SVR)

In this section the BHA-SVR model will be described to determine the best hyperparameters that can be used in the SVR method. The procedure proposed in this paper is the process of initializing hyperparameters using a BHA and then using these hyperparameters in the SVR technique. As a random set of stars is determined in the search area and within a specified range, the objective function is evaluated by determining the fitness value of the set and selecting the best star in this set, which has the lowest cost and best fitness value, to be the BH and the remnant of the usual stars, and then the BH absorption process begins, where a new star is born after the death of a star, until a new star arrives at a lower cost and better fitness value, so that the BH moves to the location of that star, to start the process again. Anis range of values are specified for each of the penalty parameter C, the Absolon parameter ε and the parameter of the Kernel function σ, where the Radial basis function will be used. Then these parameters are entered in a frequency within a predetermined maximum. The best combination of these hyperparameters is obtained for use in SVR training to obtain an accurate prediction. This procedure can be described like the steps below:

1. The data is initially divided into the training and test group.
2. Create a set of stars at random locations in the search space and then evaluate the objective function for each star, so that the BH is chosen based on N stars, which is the size of the population, the star that has the best fitness value among the randomly generated stars.
3. Determine the maximum iteration. tmax=100

4. Start with the value of the penalty parameter C within a range of values, a maximum and a minimum. from distribution as uniform [0,1]
5. Starting with the value of the epsilon parameter ε within a range of values, a maximum, and a minimum. from distribution as uniform [0,1]
6. Starting with the value of the Kernel function parameter σ within a range of values, maximum and minimum. from distribution as uniform [0,1]
7. Set the time taken to select the best combination of hyperparameters using the BHA.
8. Start applying the BHA and calculate the fitness value of each star.
9. Satisfying the algorithm after reaching the upper limit of the number of iterations or good value of fitness.
10. Determine the best combination of hyperparameters, and use this combination to train SVR.
11. Use the model for prediction.

5. Datasets

Two sets of data were used, and both the grid search algorithm and the BHA were applied with the same settings for the values of the hyperparameters, and after obtaining the best combination of hyperparameters from each algorithm, they were used in the SVR method for each algorithm with computing the square root of the mean squares error of the training and test data. The time taken was calculated for each algorithm. This process was performed 10 times, each time with 100 iterations were used to test the proposed method.

Table 1: Dataset description

Datasets	*Samples*	*Features*
Dataset I	*60*	*635*
Dataset II	*108*	*2436*

6. Results and discussion

In this section, a model was trained using the SVR method when entering a combination of hyperparameters that were obtained using a BHA, and then this performance was compared with the performance of the same method, but using another combination of hyperparameters, which was obtained using the grid search algorithm. The SVR method and the two algorithms (grid search and the black hole) were applied to two data sets, and the criterion square root of the mean squares error was used for the test and training data to compare the two algorithms. The time taken for each algorithm was also calculated for a comparison between them.

From the test group in Table 2, the proposed BHA produces high predictive power and performance compared to the grid search algorithm. For example, in Dataset I, was 0.63609296 which was better than 0.70210495, and in Dataset II, 32.6318118 was better than 35.175248.

In order to compare the two algorithms accurately, Table 2 shows the average of ten times that were performed for each dataset.

Table 2: Ten times (on average) based on training and test datasets

Datasets	*Algorithm*	*RMSE Train*	*RMSE Test*
Datasets I	*Grid*	*0.70210495 ± 0.1466*	*1.2083129 ± 0.0716*
	BHA	*0.63609296 ± 0.1491*	*1.1889001 ± 0.0724*
Datasets II	*Grid*	*35.175248 ± 14.4076*	*54.16619 ± 5.4607*
	BHA	*32.6318118 ± 19.1028*	*54.113825 ± 5.4783*

From Table 2, we notice that the average value of the square root of the mean squares error for the (RMSE Train and Test) of the BHA is less than the grid search algorithm for the two datasets (Melting and Flu).

Furthermore, the superiority of the black hole algorithm in terms, the computational time for choosing the best combination of the hyperparameters by comparison to the grid search algorithm, which is shown in Figure 1.

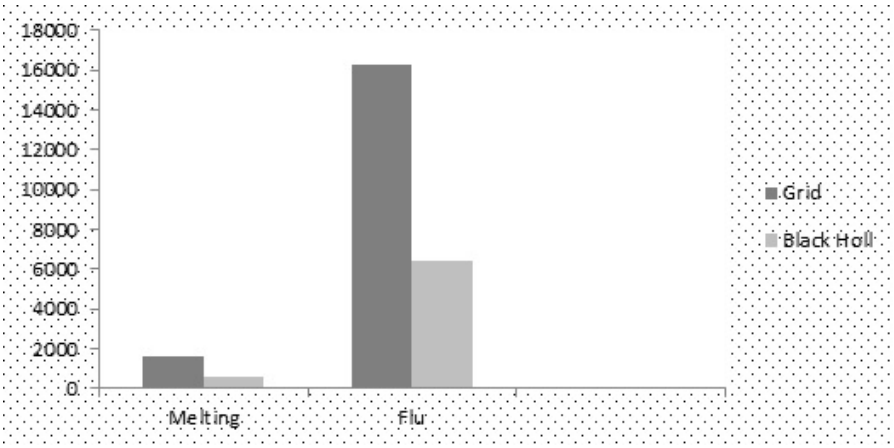

Figure 1: The time taken in seconds to select the best combination of hyperparameters

It is also possible to notice the big difference in the average time shown in Table 3, which shows the preference of the computational time for the black hole algorithm compared to the grid search algorithm.

Table 3: Average time taken for the two algorithms, per minute.

Datasets	*Algorithm*	*The average time in second*
Dataset I	*Grid*	*16286.735*
	BHA	*6448.618*
Dataset II	*Grid*	*1615.695*
	BHA	*594.036*

Also, a significant test of the time taken to find the best combination of the Hyperparameters of the two algorithms was performed. The test was significant and the result can be seen in Table 4.

Table 4: Time t-test for two algorithms

Datasets	*t*	*p-Value*
Melting	*11.026*	*0.000*
Flu	*10.34*	*0.000*

7. Conclusion

In this paper, a new mechanism was developed to choose the best combination of hyperparameters, by relying on the BHA instead of using the grid search or another nature-inspired algorithm. As the empirical results and statistical analysis on two datasets showed a superiority of the proposed algorithm over a remnant of the methods and other algorithms in terms of high prediction performance and runtime. It is clear that the proposed algorithm is an effective and promising method for optimizing hyperparameters and increasing predictive power, and this effective frame can be applied to many other real apps.

Acknowledgments

This work was supported by the Mosul University, College of Computer Sciences and Mathematics, Republic of Iraq.

References

[1] A.A. Akinpelu, M.E. Ali, T.O. Owolabi, M.R. Johan, R. Saidur, S.Q. Olatunji and Z. Chowdbury, *A support vector regression model for the prediction of total polyaromatic hydrocarbons in soil: An artificial intelligent system for mapping environmental pollution*, Neural Comput. Appl. 32(18) (2020) 14899–14908.

[2] N.A. Al-Thanoon, Q.S. Qasim and Z.Y. Algamal, *Selection of tuning parameter in L1-support vector machine via. particle swarm optimization method*, J. Engin. Appl. Sci. 15(1) (2019) 310–318.

[3] Z. Qasim, M. Lee and H. Ali, *Improving grasshopper optimization algorithm for hyperparameters estimation and feature selection in support vector regression*, Chem. Intell. Lab. Syst. 208 (2021) 104196.

[4] H. Aslani, M. Yaghoobi and M. R. Akbarzadeh, *Chaotic inertia weight in black hole algorithm for function optimization*, Int. Cong. Technol. Commun. Knowledge 2015.

[5] K.Y. Chen and C.H. Wang, *Support vector regression with genetic algorithms in forecasting tourism demand*, Tourism Manag. 28(1) (2007) 215–226.

[6] V. Cherkasskyand and Y. Ma, *Selection of Meta-parameters for Support Vector Regression*,Artif. Neural Networks, Springer, Berlin, Heidelberg, 2002, pp. 687–693.

[7] V. Cherkassky and Y. Ma, *Practical selection of SVM parameters and noise estimation for SVM regression*, Neural Networks 17(1) (2004) 113–126.

[8] J.S. Chou and A.D. Pham, *Nature-inspired metaheuristic optimization in least squares support vector regression for obtaining bridge scour information*, Inf. Sci. 399 (2017) 64–80.

[9] C.C. Chuang and Z.J. Lee, *Hybrid robust support vector machines for regression with outliers*, Appl. Soft Comput. 11(1) (2011) 64–72.

[10] M. Eskandarzadehalamdary, B. Masoumi and O. Sojodishijani, *A new hybrid algorithm based on black hole optimization and bisecting k-means for cluster analysis*, 22nd Iran. Conf. Electric. Engin. 2014, pp. 1075–1079.

[11] M. Farahmandian and A. Hatamlou, *Solving optimization problems using black hole algorithm*, J. Adv. Comput. Sci. Technol. 4(1) (2015) 68.

[12] R. Fernandez, *Predicting time series with a local support vector regression machine*, ACAI 99, 1999.

[13] W. Gao, X. Wang, S. Dai and D. Chen, *Study on stability of high embankment slope based on black hole algorithm*, Envir. Earth Sci. 75(20) (2016) 1381.

[14] A. Hatamlou, *Black hole: A new heuristic optimization approach for data clustering*, Inf. Sci. 222 (2013) 175–184.

[15] H. Kaneko and K. Funatsu,*Fast optimization of hyperparameters for support vector regression models with highly predictive ability*, Chemometrics and Intelligent Laboratory Systems, 142 (2015) 64-69.

[16] R. Laref, E. Losson, A. Sava and M. Siadat, *On the optimization of the support vector machine regression hyperparameters setting for gas sensors array applications*, Chem. Intell. Lab. Syst. 184 (2019) 22–27.

[17] S. Fang and X. Liu, *Parameter optimization of support vector regression based on sine cosine algorithm*, Expert Syst. Appl. 91 (2018) 63–77.

[18] M. Nait Amar and N. Zeraibi,*Application of hybrid support vector regression artificial bee colony for prediction of MMP in CO2-EOR process*, Petroleum, 6(4) (2018) 415–422.

[19] E. Pashaei, M. Ozen and N. Aydin, *An application of black hole algorithm and decision tree for medical problem*, IEEE 15th Int. Conf. Bioinf. Bioengin. 2015, pp. 1–6.

[20] O.S. Qasim, K.A. Abed and A.F. Qasim, *Optimal parameters for nonlinear Hirota-Satsuma coupled KdV system by using hybrid firefly algorithm with modified Adomian decomposition*, J. Math. Fund. Sci. 52(3) 2020 339–352.
[21] B. Schölkopf, A.J. Smola, R.C. Williamson and P.L. Bartlett, *New support vector algorithms*, Neural Comput. 12(5) (2000) 1207–1245.
[22] A. Smola, N. Murata, B. Schölkopf and K.R. Müller, *Asymptotically optimal choice of ε-loss for support vector machines*, Int. Conf. Artif. Neural Networks, 1998, pp. 105–110.
[23] D.D. Warnana, *Black hole algorithm for determining model parameter in self-potential data*, J. Appl. Geophys. 148 (2018) 189–200.
[24] B. Üstün, M.J. Melssen, M. Oudenhuijzen and L.M. Buydens, *Determination of optimal support vector regression parameters by genetic algorithms and simplex optimization*, Anal. Chim. Acta 544(1-2) (2005) 292–305.
[25] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer, new york, 1995
[26] V. Vapnik, S.E. Golowich and A. Smola, *Support vector method for function approximation, regression estimation, and signal processing*, Adv. Neural Inf. Process. Syst. (1997) 281–287.
[27] S. Xu, B. Lu, M. Baldea, T.F. Edgar and M. Nixon, *An improved variable selection method for support vector regression in NIR spectral modeling*, J. Process Cont. 67 (2018) 83–93.
[28] Y. P. Zhao and J. G. Sun, *Robust truncated support vector regression*, Expert Syst. Appl. 37(7) (2010) 5126–5133.