Indonesian Journal of Electrical Engineering and Computer Science
Vol. 28, No. 2, November 2022, pp. 1088~1095
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v28.i2.pp1088-1095

Determination of support vector regression parameters using African buffalo optimization algorithm

Inusa Sani Maijama'a[1,2], Yuhanis Yusof[3], Mohamad Farhan Mohsin[3]
[1]Department of Computer Science, School of Computing, Universiti Utara Malaysia, Kedah, Malaysia
[2]Department of Computer Science, Hussaini Adamu Federal Polytechnic, Kazaure, Nigeria
[3]Department of Computer Science, School of Computing, Universiti Utara Malaysia, Kedah, Malaysia

Article Info

Article history:

Received Jul 18, 2021
Revised Aug 1, 2022
Accepted Aug 26, 2022

Keywords:

African buffalo optimisation
Computational intelligence
Data mining
Machine learning
Parameter optimisation
Support vector regression

ABSTRACT

The use of support vector regression (SVR) for regression tasks has been on increase over the past few years. Unfortunately, the practical application of SVR for regression task is limited due to its dependence on proper setting of its hyper-parameters and associated kernel parameter. Therefore, it become imperative to device a reliable and fast mechanism of determining the value of these parameters that could guarantee lowest generalization error. This paper presents SVR parameter optimization approaches using African buffalo optimisation (ABO) algorithm, i.e. SVR-ABO. The SVR parameters are optimized by using African buffalo optimisation algorithm. Results obtained from several experiments performed has shown that the proposed ABO algorithm has the capability of determining SVR hyper-parameters which most of time has to be done through estimation.

Corresponding Author:

Inusa Sani Maijama'a
School of Computing, University Utara Malaysia
Sintok, 06010, Bukit Kayu Hitam, Kedah, Malaysia
Email: Inusa_sani@ahsgs.uum.edu.my

1. INTRODUCTION

Support vector machines (SVM) [1]–[7] as a statistical learning theory based machine learning technique introduced [6], [8] has gained popularity in both classification and regression tasks due to its various associated prominent abilities that led to a promising generalization performance [6], [9]. Some of the most distinguished features of SVR include its ability to model non-linear relationship among data samples, non-dependence on input space dimensionality for its generalization and always resulting to global solution due to its quadratic formulation approach. Unlike ANN, SVR overcome the overfitting problem through adoption of structural risk minimization (SRM) approach that focus on minimizing the generalization error instead of minimizing training errors [10].

However, despite the mentioned advantages of SVR, the performance of the technique depends on appropriate settings three free parameters namely regularization parameter C, tube size ε and kernel parameter γ [11], [12]. Presently, there is no definitive method of selecting these free parameters that can be found in literature. Different approaches were proposed, however most of these approaches are dependent on user's domain knowledge and expertise or through trial and error while some are even contradictory [13]. Hence, there is lack of certainty about optimality of parameters' values obtained. On the other hand, the problem associated with SVR parameters selection become exacerbated due to the fact that these three parameters have to be simultaneously determined for SVR to achieve good generalization ability. Based on mentioned reasons, many researchers opt for grid search or cross-validation (CV) methods as an alternative method for determining optimal SVR parameters. However, grid search and CV methods have been reported to be computationally expensive, time-consuming and mostly reported high error rate [14]–[16]. Hence, grid search and CV methods are considered not suitable choice for determining SVR parameters.

This led researchers to explore swarm intelligence (SI) methods to optimize optimal parameters of SVR algorithm as can be established in literature [2], [3], [5], [17]–[21]. SI algorithms like PSO has been widely used for parameter optimization for several algorithms including SVR. The results have shown to be providing promising results. Recently, a new SI-based optimisation algorithm namely African buffalo optimisation (ABO) has been introduced into research community by Odili and Noraziah [22]. ABO algorithm has been used with many algorithms as an optimizing agent in various areas such as travelling salesman problem (TSP) [23], selection of biodiversity conservation area with a given constraint [24] and parameter tuning of PID controller [25]. In all the mentioned cases, ABO reported success against compared similar algorithms due to its fast convergence, ability to track the best position, and speed of each buffalo as well as the movement of best buffalo towards better exploration [26]. However, the performance of ABO in optimizing three mutual parameters simultaneously as in the case of SVR has not been recorded in the literature. Hence, this study proposed to hybridize SVR with ABO as an optimization algorithm to determine optimal parameters of SVR.

The main objective of this paper is to use ABO algorithm to auto-determine optimal values of SVR using ABO algorithm. The developed hybrid algorithm reported higher forecasting accuracy as compared with standard method of opting for the default SVR parameters. The remaining sections of this paper presented various processes involved towards the development of the SVR-ABO algorithm.

2. THEORETICAL CONCEPT

2.1. Support vector regression (SVR)

This section present the theory behind SVR equations as given by [27], [28]. Given a regression problem as (1).

$$D = \{(p_1, q_1)(p_2, q_2)\ldots(p_n, q_n)\} \tag{1}$$

Where *D* represent the dataset, $p \in P \subset \mathbb{R}^n$ are the training inputs and $q \in Q \subset \mathbb{R}$ are the training outputs. The main objective is to determine a given function that can approximate the existing relationship between the inputs variable(s) represented as *p* and the associated target variable represented as *q*. The function later can be used to infer new value of target variable *q* in the future, given new input data value(s). For any regression function *f(p)*, there is a loss function *L* that determine the amount of deviation by the function's output as from the actual value. In this paper we adopted the commonly use loss function that was proposed by Gunn [29] and formulated as (2).

$$L(p; g(q)) = \begin{cases} 0 \ \ if |p - g(q)| \leq \ \varepsilon \\ |p - g(q)| - \ \varepsilon \ \text{otherwise} \end{cases} \tag{2}$$

Assuming the linear function *g* above is represented as (3):

$$f(q) = w.p + b \tag{3}$$

where *p* is an element of vector *P* in input space *P*, where ***w*** is the weight vector, *b* is the bias and *w.p* represent a dot product operation of vector *w* and *p*. The primary objective of the formulated regression function is to fit the data points with a function that is flat. In the case of (3), the flatness can be achieved by making the value of *w* to be as small as possible. One way to achieve such flatness is by minimizing the norm i.e. w^2. By so doing, the regression problem can be formulated as an optimization problem in convex form as follows:

$$minimise \quad \frac{1}{2}\, \boldsymbol{w^2} \tag{4}$$

$$subject\ to\colon \begin{cases} q_i - (w^T\Theta(p_i) + b) \leq \varepsilon \\ q_i - (w^T\Theta(p_i) + b) \ \geq \ \varepsilon \end{cases} \tag{5}$$

Hence, the (4) can be represented as (6), (7).

$$minimize\ \frac{1}{2}\ \|\boldsymbol{\omega}\|^2 + \boldsymbol{C}\sum_{i=1}^{N}(\xi_i + \ \xi_i^*) \tag{6}$$

$$\begin{cases} ((w^T \Theta(p_i) + b) - y_i) \leq \varepsilon + \xi_i \\ y_i - (w^T \Theta(p_i) + b) \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \tag{7}$$

The value of the constant *C* is used to determine the amount of flatness that can be traded-off as the result of allowable deviations that are larger than the tolerable value of ε. This trade-off value corresponds to the value of ε-insensitive loss function that was described before which is represented in (6) by $(\xi_i + \xi_i^*)$ which are also called empirical terms. The empirical terms formed the ε-tube which is the acceptable maximum deviation. All data points that fall within the range of allowable deviation have zero valued coefficient, hence have non-significant contribution value.

Upon application of Krush-Kuhn_Tucker as a means for finding Lagrangian solution, all products of the dual variables vanished at solution point, thus resulting (2) into (8).

$$f(x) = \sum_{i=1}^{m} (\alpha_j^* - \alpha_j) K(p_i, p_j) + b \tag{8}$$

where $\boldsymbol{p_i}, \boldsymbol{p_j}$ correspond to the support vectors.

Various kernel functions can be found in literature, however, radial basis function has been established to be most appropriate for regression task [4], [27] which is represented as (9):

$$K(x_i, x_j) = e^{-\gamma(\|x_i, x_j\|^2)} \tag{9}$$

where γ represents a positive non-zero kernel parameter. Hence, (5) can be reduced to (10).

$$f(x) = \sum_{i=1}^{m} (\alpha_i^+ - \alpha_i^-) K(x_i, x_j) + b \tag{10}$$

The performance of SVR algorithm is dependent on appropriate settings of its hyper-parameters viz regularization constant *C,* error-sensitivity parameter ε, and kernel parameter γ [16]. The regularization parameter determines the tolerable bound of deviations that are larger than defined value of error sensitivity parameter and the degree of model complexity. Inappropriate selection of *C* could lead to an imbalance between model complexity minimization (MCM) and empirical risk minimization (ERM). The error-sensitivity parameter (ε) defines allowable ε-insensitive zone based on number of support vectors (SV). Large error-sensitivity value results into accommodating significant number of data points into insensitive zone, hence undesirable regression estimates due to fewer support vector points. The kernel parameter (γ) is used to determine the degree of wideness allowable by RBF kernel. Large value of kernel parameter results into a non-flexible function, unsuitable for complex function approximation. While smaller value of kernel parameter produces over-flexible function that could results into overfitting.

2.2. African buffalo optimization algorithm

African buffalo optimisation (ABO) algorithm as swarm-intelligence based algorithm that is characterized by a high speed convergence feature developed by Odili *et al.* [30]. The algorithm was developed based on the foraging and herd defending behaviour exhibited by the wild African buffaloes [30], [31]. These wild animal exhibit exceptional organizing behaviour of which ABO optimisation algorithm was model based upon [23], [32]. In the ABO algorithm as presented in (11) and (12), the w_i, represents the "*waaa*' sound, m_i represents the "*maaa*" sound, while l_1 and l_2 represents its two learning parameters. Other distinct parameters of the algorithm are global best (bg_{max}), the personal best ($bp_{max(k)}$) positions. The basic ABO algorithm is controlled by democratic and location update equations represented by (11) and (12) respectively. The algorithm operates by deducting the "waaa" value (w_k) from both global best (bg_{max}) and the personal best ($bp_{max(k)}$) which are enhanced by learning parameters in order to direct the herd to either explore the search space for better pasture or to retain their present position and continue grazing. The detailed of the ABO algorithm operation stages is depicted in (11) and (12) respectively [30].

$$m_{k+1} = m_k + l_1(bg_{max} - w_k) + l_2(bp_{\max(k)} - w_k) \quad (11)$$

$$w_{k+1} = \frac{(w_k + m_k)}{\lambda} \quad (12)$$

3. METHOD

The methodology involved in developing the proposed hybrid technique composed of different stages ranging from procedure involved, type of dataset used to build and test the model. The proposed hybrid model was developed using python programming language with numpy and pandas libraries. Scikit-learn version of LibSVM implementation was used as to build the SVR model. The detailed methodology is as describe in the following sub-sections.

3.1. Dataset description

Four (4) multivariate historical datasets were used to test the accuracy of the developed hybrid algorithm. The datasets were obtained using Quandl API. In order to rigorously test the developed model, we used four datasets of varying properties that include different number of parameters and number of instances. Google and Amazon dataset have twelve (12) parameters each while Tesla and Yahoo have six (6) parameters each. In terms of size, we opt to use the Tesla dataset which has only 251 instances in order to test the performance of the hybrid model on small dataset. Description of the datasets is as presented in Table 1.

Table 1. Properties of the datasets

S No	Dataset Name	Duration	No of Instances	No. of Variables
1.	Google Stock	21/03/2018-19/08/2004	3,424	12
2.	Tesla Stock	18/03/2019-23/03/2018	251	6
3.	Yahoo stock	11/12/2019-03/01/2012	2,003	6
4.	Amazon Inc	16/05/1997-21/03/2018	5,248	12

The Tesla Inc dataset consists of two hundred and fifty-one (251) instances of daily recordings. The Yahoo Inc dataset consists of two thousand and three (2003) instances of daily recordings. The Amazon Inc dataset consists of Five thousand two hundred and forty-eight (5,248) number of instances while the Google Inc consists of Three four hundred and twenty-four (3,424).

Each of the datasets was portioned into three (3) parts viz; 70% for training, 15% for validation during training and remaining 15% for the actual testing the model. After the partitioning, each of the dataset was scaled to values between {1 and -1} in order to remove the differences in magnitude of features to avoid influence of features with higher magnitude. The training and validation datasets were scaled separately from testing dataset in order to avoid data leakage between the model training and testing phases.

3.2. Evaluation metrics

In order to evaluate the performance of the developed hybrid model. We use mean absolute percentage (MAPE) and root mean squared error (RMSE) as the two regression evaluation metrics whose sole aim is to test the level of accuracy of all developed models. The aim of these metrics is to obtain a small value; the smaller the MAPE or RMSE, the better the performance of the developed model is. These performance metrices are mathematically represented as in (15) and (16) respectively.

$$MAPE = \frac{1}{N}\left[\sum_{n=1}^{N} \left| \frac{y_n - \acute{y}_n)}{y_n} \right|\right] * 100\% \quad (13)$$

Where N represents number of observations, y_n and $\acute{y}_n$ represent the n_{th} observed and forecast values respectively.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(a_i - f_i)^2}{N}} \quad (14)$$

3.3. The Hybrid SVR-ABO algorithm

As already pointed out in previous section about the difficulty in determining the three (3) parameters of SVR, in our proposed hybrid algorithm we hybridized the SVR algorithm with ABO algorithm. The ABO is used iteratively to find optimal values that correspond to the SVR algorithm's parameters. The process started by initializing the buffalo population in search space using Mersenne Twister algorithm for better randomization. The fitness of each buffalo was determined based on its initial position in the search space. The buffalo that returns the least prediction error based RMSE error metric, is considered as the best for that iteration. The overall best buffalo among the best buffaloes produce at each iteration is considered as the global best. The training process was conducted based on population size of Two hundred (200) and maximum number of iterations of one thousand (1,000).

Upon reaching the termination criteria, the values of the position of the overall best buffalo are considered the required optimal hyper-parameters for the SVR model. Then, the SVR with optimized parameters was used to forecast the test dataset. Algorithm 1 illustrate the SVR-ABO forecasting process.

Algorithm 1: **SVR-ABO Algorithm**

Input: *Training Data, P D, Max_I, l1, l2*
/ P= Number of individual Buffaloes (population size), D= Problem dimension (SVR control parameters), Max_I = Maximum number of Iterations, l1= Cognitive Learning parameter, l2= Social Learning parameter */*
Output: *Optimal values for SVR (**C**, γ and **ε**) as Global best buffalo position*
For buffalo i = 0 to P **do**:
    Random initialisation of buffalo position vector m_k with three (3) values based on [**C**, γ and **ε**] ranges
    Random initialisation of each buffalo movement vector w_k
End For
Initialise t = 1
While (t ≠ *Max_I)* **do**:
    For each buffalo i **do**:
        Calculate fitness_value using SVR regressor
        If buffalo's fitness_value is better than $bp_{\max(k)}$
            Set $bp_{\max(k)}$ = buffalo's current fitness
        End If
    End For
    Set bg_{max} = Best previous buffalo's fitness_value
/* Updating each buffalo's movement and position */
    For buffalo i = 0 to P ***do:***
        For dimension d = 0 to D ***do:***

$$m_{id}^{t+1} = m_{id}^{t} + l_1(bg_{max} - w_{id}^{t}) + l_2(bp_i^t - w_{id}^{t})$$

$$w_{id}^{t+1} = \frac{(w_{id}^{t} + m_{id}^{t+1})}{\lambda}$$

        End For
    End For
    Set t = t + 1
End While
Evaluate the solution on testing set
Result: The forecasting values and performance measurement on the testing set

4. RESULT AND DISCUSSION

The model developed based on the hybrid algorithm was tested using four (4) different stock market historical datasets. The test datasets as described in methodology section were used to test the performance of the developed model based on RMSE and MAPE as two statistical evaluation metrics. The result obtained from the developed model was compared with results obtained from classical SVR, PSO and GA algorithms all on default parameters settings. The findings obtained has shown that ABO algorithm can find optimal parameters for SVR algorithm better than the remaining benchmarked algorithms.

The following analysis, as presented in Table 2 shows the comparison performance of the proposed SVR-ABO algorithm against classical SVR, SVR-PSO and SVR-GA algorithms based on MAPE metric. The performance of the SVR-ABO algorithm on Amazon dataset shows that it is able to achieves 96.67% accuracy based MAPE, while classical SVR, SVR-PSO and SVR-GA recorded 96.09%, 96.48%, and 96.45% respectively. These algorithms recorded RMSE values of 36.4381 for classical SVR, 34.5642 for SVR-GA, 33.6503 for SVR-ABO, and 33,9624 for SVR-PSO.

On Google dataset, the algorithms were able to record RMSE values of 35.3514, 48.9568, 33.2592 and 33.5197 for classical SVR, SVR-GA, SVR-ABO and SVR-PSO respectively. While on MAPE metric, the algorithms performance shows that the SVR-ABO algorithm recorded accuracy of 96.65%, while the classical SVR, SVR-PSO, and SVR-GA algorithms were able to record accuracy of 96.28%, 96.50%, and 95.45% respectively.

The performance of the algorithms on Tesla test data has shown that SVR-ABO is able to achieve higher performance with accuracy of 96.83%, while SVR-PSO, SVR-GA and classical SVR were able to achieve accuracy of 96.40%, 96.82%, and 95.81%. These algorithms recorded RMSE values of 35.7663 for classical SVR, 34.7329 for SVR-GA, 34.7005 for SVR-ABO, and 35.0092 for SVR-PSO.

Lastly, on Yahoo test data, the algorithms recorded RMSE values of 35.9359, 33.9286, 32.2832 and 32.9706 for classical SVR, SVR-GA, SVR-ABO and SVR-PSO respectively. While on MAPE metric, the algorithms performance shows that the SVR-ABO algorithm recorded accuracy of 96.21%, while the classical SVR, SVR-PSO, and SVR-GA algorithms were able to record accuracy of 95.75%, 96.04%, and 96.02% respectively.

Table 2. Evaluation of SVR and SVR-ABO over four (4) datasets based on RMSE and MAPE

		RMSE				MAPE			
		SVR	SVR-GA	SVR-ABO	SVR-PSO	SVR	SVR-GA	SVR-ABO	SVR-PSO
Datasets	Amazon	36.4381	34.5642	**33.6503**	33.9624	3.9074	3.5491	**3.3350**	3.5164
	Google	35.3514	48.9568	**33.2592**	33.5197	3.7185	4.5537	**3.4799**	3.5019
	Tesla	35.7663	34.7329	**34.7005**	35.0092	4.1859	3.17601	**3.1741**	3.6050
	Yahoo	35.9359	33.9286	**32.2832**	32.9706	4.2337	3.9786	**3.7912**	3.9628

Figures 1-4 shows the visual performance of the forecasting result obtained by the proposed algorithm against other benchmarked algorithms on the testing data reserved for evaluation purpose. Figure 1 shows the predicted values against the actual values on Amazon dataset for a period of one month. Figure 2 shows the comparison performance of the algorithms on Google test. Figure 3 shows the visual performance of the algorithms on Tesla test data over a period of last three (3) months. Lastly, Figure 4 shows comparison performance between the developed hybrid algorithm and other algorithms on Yahoo test data.

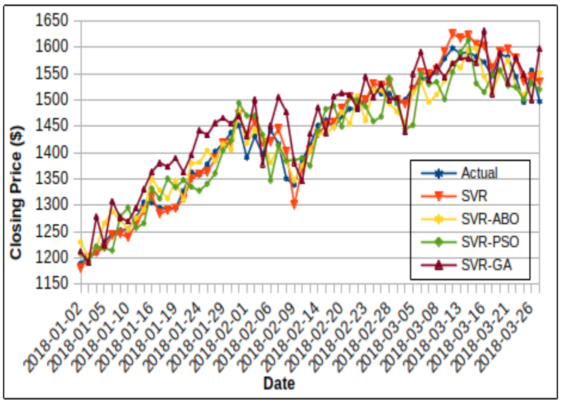

Figure 1. SVR-ABO performance on Amazon dataset

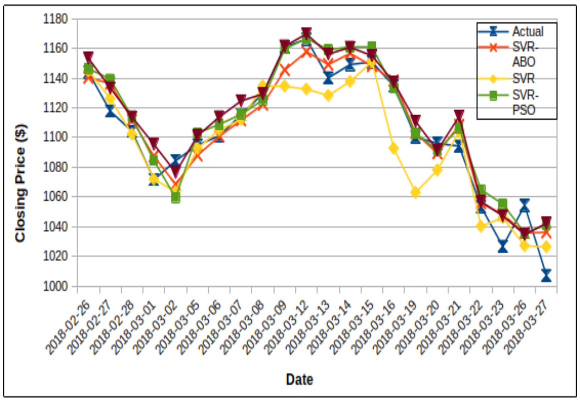

Figure 2. SVR-ABO performance on Google dataset

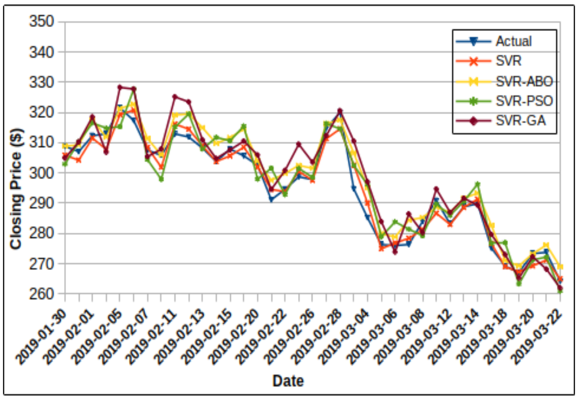

Figure 3. SVR-ABO performance on Tesla test data

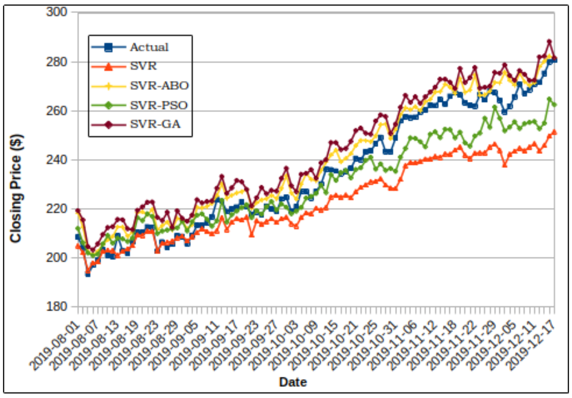

Figure 4. SVR-ABO performance on Yahoo dataset

5. CONCLUSION AND FUTURE WORK

In this paper, a new method of determining the optimal values of SVR algorithm based on African buffalo optimisation (ABO) algorithm has been proposed. The proposed method has shown that it can achieve higher accuracy in terms of generalization than classical SVR with default settings of parameters. The proposed method has been tested using stock price data of some selected companies. The time series datasets were chosen due to their inherent nature of non-linearity and the dependence of the closing stock price of each day on several variables (multivariate). The key feature of the proposed approach is its ability of exploiting the power of swarm intelligence method to auto-determine optimal parameters values of classical SVR algorithm. Hence, the method has the capability of increasing generalization ability of SVR thereby increases model accuracy. The future work will explore the performance of the developed algorithm on other dataset in different domains.

REFERENCES

[1] D. Gupta and U. Gupta, "On robust asymmetric Lagrangian ν-twin support vector regression using pinball loss function," *Appl. Soft Comput.*, vol. 102, 2021, doi: 10.1016/j.asoc.2021.107099.

[2] O. Chamorro-atalaya, G. Morales-romero, Y. Meza-chaupis, and O. Chamorro-atalaya, "Quadratic vector support machine algorithm, applied to prediction of university student satisfaction," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 27, no. 1, pp. 139–148, 2022, doi: 10.11591/ijeecs.v27.i1.pp139-148.

[3] I. S. Maijama'a, Y. Yusof, and M. F. Mohsin, "Short-term load forecasting using supprt vector regression with African-buffalo optimization algorithm," *J. Theor. Appl. Inf. Technol.*, vol. 99, no. 8, pp. 1706–1718, 2021.

[4] E. A. Gheni and Z. M. Algelal, "Human face recognition methods based on principle component analysis (PCA), wavelet and support vector machine (SVM): a comparative study," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 20, no. 2, pp. 991–999, 2020, doi: 10.11591/ijeecs.v20.i2.pp991-999.

[5] E. Hazir, T. Ozcan, and K. H. Koç, "Prediction of adhesion strength using extreme learning machine and support vector regression optimized with genetic algorithm," *Arab. J. Sci. Eng.*, vol. 45, no. 8, pp. 6985–7004, 2020, doi: 10.1007/s13369-020-04625-0.

[6] M. H. Jopri, A. R. Abdullah, J. Too, T. Sutikno, S. Nikolovski, and M. Manap, "Support-vector machine and Naïve Bayes based diagnostic analytic of harmonic source identification," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 20, no. 1, pp. 1–8, 2020, doi: 10.11591/ijeecs.v20.i1.pp1-8.

[7] H. Shi, H. Xiao, J. Zhou, N. Li, and H. Zhou, "Radial basis function kernel parameter optimization algorithm in support vector machine based on segmented dichotomy," *2018 5th Int. Conf. Syst. Informatics, ICSAI 2018*, no. Icsai, 2019, pp. 383–388, doi: 10.1109/ICSAI.2018.8599461.

[8] D. Saini and A. Saxena, "Electric price forecast using interbreed approach of linear regression and SVM," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 2, no. 3, pp. 537–544, 2016, doi: 10.11591/ijeecs.v2.i3.pp537-544.

[9] K. Kumar and D. P. Gandhmal, "An intelligent indian stock market forecasting system using LSTM deep learning," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 21, no. 2, pp. 1082–1089, 2020, doi: 10.11591/ijeecs.v21.i2.pp1082-1089.

[10] M. S. Al-Musaylh, R. C. Deo, J. F. Adamowski, and Y. Li, "Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia," *Adv. Eng. Informatics*, vol. 35, no. April 2017, pp. 1–16, 2018, doi: 10.1016/j.aei.2017.11.002.

[11] M. Sarhani, A. E. Afia, and A. El Afia, "Generalization enhancement of support vector regression in electric load forecasting with model selection," in *Proceedings of the International Conference on Learning and Optimization Algorithms: Theory and Applications - LOPAL '18*, 2018, pp. 1–6. doi: 10.1145/3230905.3230947.

[12] S. Zhang, F. Kuang, and R. Hu, "Support vector regression with multi-strategy artificial bee colony algorithm for annual electric load forecasting," *Adv. Intell. Syst. Comput.*, vol. 891, pp. 576–585, 2019, doi: 10.1007/978-3-319-68527-4.

[13] Y. Ren and G. Bai, "Determination of optimal SVM parameters by using GA/PSO," *J. Comput.*, vol. 5, no. 8, pp. 1160–1168, 2010, doi: 10.4304/jcp.5.8.1160-1168.

[14] Q. Bing, D. Qu, X. Chen, F. Pan, and J. Wei, "Short-term traffic flow forecasting method based on LSSVM model optimized by GA-PSO hybrid algorithm," *Discret. Dyn. Nat. Soc.*, vol. 2018, p. 10, 2018, doi: 10.1155/2018/3093596.

[15] J. X. Che, Y. L. Yang, L. Li, Y. Y. Li, and S. L. Zhu, "A modified support vector regression: integrated selection of training subset and model," *J. Appl. Soft Comput.*, vol. 53, pp. 308–322, 2017, doi: 10.1016/j.asoc.2016.12.053.

[16] Y. Yusof and Z. Mustaffa, "A review on optimization of least squares support vector machine for time series forecasting," *Int. J. Artif. Intell. Appl.*, vol. 7, no. 2, pp. 35–49, 2016, doi: 10.5121/ijaia.2016.7203.

[17] D. Karthika and K. Karthikeyan, "Estimation of electrical energy consumption in Tamil Nadu using univariate time-series analysis," *Ann. Optim. Theory Pract.*, vol. 4, no. 2, pp. 31–37, 2021, doi: 10.22121/AOTP.2021.292718.1073.

[18] M. O. Mustafa, "Optimal parameter values of pid controller for DC motor based on modified particle swarm optimization with adaptive inertia weight," *Eastern-European J. Enterp. Technol.*, vol. 1, pp. 35–45, 2021, doi: 10.15587/1729-4061.2021.225383.

[19] E. Dodangeh *et al.*, "Novel hybrid intelligence models for flood-susceptibility prediction: Meta optimization of the GMDH and SVR models with the genetic algorithm and harmony search," *J. Hydrol.*, vol. 590, no. July, p. 125423, 2020, doi: 10.1016/j.jhydrol.2020.125423.

[20] X. Zhang, H. Huang, and X. Zhang, "Applying cuckoo search algorithm to solve fractional differential equation based on cubic spline function," *IAENG Int. J. Appl. Math.*, vol. 50, no. 1, pp. 31–42, 2020.

[21] S. Amosedinakaran, K. Mala, A. Bhuvanesh, and P. Marishkumar, "Electricity demand forecasting using differential evolution algorithm for Tamil Nadu," *Proc. 4th Int. Conf. Trends Electron. Informatics, ICOEI 2020*, no. Icoei, 2020, pp. 143–146, doi: 10.1109/ICOEI48184.2020.9142872.

[22] J. B. Odili and A. Noraziah, "African buffalo optimization for global optimization," *Curr. Sci.*, vol. 114, no. 3, pp. 627–636, 2018, doi: 10.18520/cs/v114/i03/627-636.

[23] J. B. Odili, M. N. M. Kahar, A. Noraziah, M. Zarina, and R. U. Haq, "Performance analyses of nature-inspired algorithms on the traveling salesman's problems for strategic management," *Intell. Autom. Soft Comput.*, vol. 8587, no. June, pp. 1–11, 2017, doi: 10.1080/10798587.2017.1334370.

[24] B. Almonacid, J. Reyes-Hagemann, J. Campos-Nazer, and J. Ramos-Aguilar, "Selecting a biodiversity conservation area with a limited budget using the binary African buffalo optimisation algorithm," *IET Softw.*, vol. 12, no. 2, pp. 96–111, 2017, doi: 10.1049/iet-sen.2016.0327.
[25] J. B. Odili, M. N. M. Kahar, and A. Noraziah, "Parameters-tuning of PID controller for automatic voltage regulators using the African buffalo optimization," *PLoS One*, vol. 12, no. 4, pp. 1–17, 2017, doi: 10.1371/journal.pone.0175901.
[26] J. Odili and M. Kahar, "Numerical function optimization solutions using the african buffalo optimization algorithm (ABO)," *Br. J. Math. Comput. Sci.*, vol. 10, no. 1, pp. 1–12, 2015, doi: 10.9734/bjmcs/2015/17145.
[27] Q. Hou, J. Zhang, L. Liu, Y. Wang, and L. Jing, "Discriminative information-based nonparallel support vector machine," *Signal Processing*, vol. 162, pp. 169–179, 2019, doi: 10.1016/j.sigpro.2019.03.026.
[28] C. Corinna and V. Vladimir, "Support-Vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995, doi: 10.1007/BF00994018.
[29] S. Gunn, "ISIS technical report: support vector machines for classification and regression," *Analyst*, vol. 135, no. 2, 1998, doi: 10.1039/b918972f.
[30] J. B. Odili, M. N. M. Kahar, and S. Anwar, "African buffalo optimization: a swarm-intelligence technique," *Procedia Comput. Sci.*, vol. 76, pp. 443–448, 2015, doi: 10.1016/j.procs.2015.12.291.
[31] O. J. Buluchi, "The dawn of meteheuristic algorithms," *Int. J. Softw. Eng. Comput. Syst.*, vol. 4, no. 2, pp. 49–61, 2018.
[32] J. B. Odili, M. N. Kahar, and A. Noraziah, "Solving Traveling salesman' s problem using african buffalo optimization, honey bee mating optimization & lin-kerninghan algorithm," *World Appl. Sci. J.*, vol. 34, no. 7, pp. 911–916, 2016, doi: 10.5829/idosi.wasj.2016.34.7.329.

BIOGRAPHIES OF AUTHORS

Inusa Sani Maijama'a is a PhD student at Data science lab, School of Computing, Universiti Utara Malaysia, Malaysia. He obtained his Masters degree from Multimedia University, Malaysia. His research interest includes Artificial Intelligence, Swarm Intelligence and Machine learning. He is a lecturer at Computer science department at Hussaini Adamu Federal Polytechnic, Kazaure, Jigawa state, Nigeria. He can be contacted at email: inusa_sani@ahsgs.uum.edu.my or inusa.sani.mjm@gmail.com.

Prof. Dr Yuhanis Yusof is a senior lecturer in School of Computing, Universiti Utara Malaysia, Malaysia. She obtained her PhD in Computer Science from Cardiff University, UK. She also holds a MSc degree in Computer Science from Universiti Sains Malaysia and Bachelor of Information Technology from Universiti Utara Malaysia. Her research interest is broadly in data analysis and management for large scale computing. This includes data mining (discovering patterns of interest from data), data warehousing, information retrieval and optimization. Currently, she is involved in several research projects involving Data Mining, Machine Learning and Swarm Computing. Her work has been applied in various domain including business management, finance, transportation, education, networking and medical. She can be contacted at email: yuhanis@uum.edu.my.

Dr. Mohamad Farhan Mohsin Mohsin is a lecturer at School of Computing, Universiti Utara Malaysia. He obtained his PhD in Data Mining and Optimization from University Kebangsaan Malaysia. He obtained his Masters degree in Information Technology from same the University. His research interest includes machine learning algorithms, Intelligent systems and Health Informatics. Dr. Farhan has published several research articles in numerous established academic journals. He can be contacted at email: farhan@uum.edu.my.