❒ 438

Acute sinusitis data classification using grey wolf optimization-based support vector machine

Ajeng Maharani Putri[1], Zuherman Rustam[2], Jacub Pandelaki[3], Ilsya Wirasati[4], Sri Hartini[5]

[1,2,4,5]Department of Mathematics, University of Indonesia, Depok, Indonesia
[3]Department of Radiology, Cipto Mangunkusumo Hospital, Jakarta, Indonesia

Article Info	ABSTRACT
	Acute sinusitis is the most common form of sinusitis, and it causes swelling and inflammation within the nose. The main thing that can causes sinusitis is probably due to viruses, and also can be caused by other factors, namely bacteria, fungi, irritation, dust, and allergens. In this research, the CT scan data attributes will be used for classification and grey wolf optimization-support vector machine (GWO-SVM) will be the machine learning technique used, where the GWO technique will be used to tuned the parameters in SVM. The performance of methods was analyzed using the python programming language with different percentages of training data, which started from 10% to 90%. The GWO-SVM method proposed provides better accuracy than using SVM without GWO. *This is an open access article under the [CC BY-SA](#) license.*

Corresponding Author:

Zuherman Rustam
Department of Mathematics
University of Indonesia
Depok 16424, Indonesia
Email: rustam@ui.ac.id

1. INTRODUCTION

Acute sinusitis is the most common form of sinusitis, and it causes swelling and inflammation within the nose [1]. The main thing that can causes sinusitis is probably due to viruses, and also can be caused by other factors, namely bacteria, fungi, irritation, dust, and allergens [2]. In diagnosing, there are several criteria that can be considered, such as nasal obstruction, facial pain, fever, and others. Some of the things mentioned earlier are included in the major criteria, while there are also minor criteria such as dental pain, headache, fatigue, cough, and others [3]. There are several methods in diagnosing sinusitis i.e. computed tomography scanning (CT scan), nasal endoscopy, nasal and sinus culture, allergy test and magnetic resonance imaging (MRI) [4]. In this research, the CT scan data attributes will be used for classification methods.

The previous research on support vector machine was proposed to process this data [4]. Also, many researches before using other methods of classification for classifying disease data, such as binary logistic regression for ovarian cancer classification [5], fast fuzzy clustering for breast cancer data [6], naive bayes classifier for predicting colon cancer [7], random forest classifier for predicting prostate cancer [8], and many more. Therefore, in this research, grey wolf optimization-support vector machine (GWO-SVM) will be the machine learning technique used, where the GWO technique will be used to tuned the parameters in SVM.

2. RESEARCH METHOD

2.1. Dataset

Acute sinusitis data from the Department of Radiology at Cipto Mangunkusumo Hospital at Jakarta in Indonesia, was used in this study. The data sets has 200 rows with two classes named acute sinusitis and non-acute sinusitis, respectively. Furthermore, this dataset contains five variables, which one of them is dependent and the rest of them are independent. The dependent variable is a class of whether the patient has acute sinusitis marked 0 or non-acute sinusitis marked 1. While, the independent variables are gender (0 for male, 1 for female), age, air cavity, and hounsfield unit (HU). Where, hounsfield unit is a relative quantitative measurement of radio density utilized by radiologists within the interpretation of computed tomography (CT) pictures [9]. These variables are used to diagnose the acute sinusitis. Table 1 shown contains three rows of data.

Table 1. Example of sinusitis data

Gender	Age	HU (Hounsfield units)	Air Cavity	Diagnosis
1	76	138	-1020	1
0	20	38	-967	0
1	57	27	-999	0

2.2. Grey wolf optimization (GWO)

Grey wolf optimization (GWO) is one of recent meta-heuristics methods which inspired by grey wolves (*Canis Lupus*), proposed by Mirjalili *et al.* [10]. The GWO algorithm simulates the leadership behavior and hunting mechanism under the grey wolf hierarchy in term of movement strategy of search agents [11]. The reason why GWO is used in this research is because based on several studies, the GWO algorithm produces better results when compared to other metaheuristic methods.

In modeling the grey wolf leadership hierarchy, alpha (α) will be considered as the most suitable solution, then beta (β) and delta (δ) will be considered as the second and third best solutions. And omega (ω) will be considered as the remaining solution. The optimization model that will be formed is a representation of the hunting behavior of the grey wolves alpha, beta, and delta, whereas omega will only following these groups [12].

Amid the hunting process, the grey wolves surround their prey. Numerically, the model of surrounding behavior is modeled by the (1) and (2) [13]:

$$\vec{D} = \vec{C}.\vec{X}_p(t) - \vec{X}(t) \qquad (1)$$

$$\vec{X}(t + 1) = \vec{X}_p(t) - \vec{A}.\vec{D} \qquad (2)$$

where
t = current iteration;
$\vec{D}$ = surrounding prey vector;
$\vec{A}, \vec{C}$ = coefficient vector;
$\vec{X}$ = a grey wolf's position vector;
$\vec{X}_p$ = the prey's position vector.

The following (3) and (4) is how the vectors $\vec{A}$ and $\vec{C}$ are determined [10]:

$$\vec{A} = 2\vec{a}.\vec{r_1} - \vec{a} \qquad (3)$$

$$\vec{C} = 2.\vec{r_2} \qquad (4)$$

where $\vec{r_1}, \vec{r_2}$ are random vectors between [0, 1] and $\vec{a}$ components are linearly reduced from 2 to 0 during the iterations [10].

To simulate the hunting behavior of grey wolves in mathematics, suppose alpha (candidate solution that the most suitable), beta and delta are the expert in knowledge pertaining to the prey's location. Hence, during the iterations, the three best solutions will be recorded, and for the next iteration calculation will be updated based on the previous three best solutions (or also called the best search agents position). The formulas defined by (5)-(7) [10], [12]:

$$\vec{D}_\alpha = |\vec{C}_1 . \vec{X}_\alpha - \vec{X}|,$$
$$\vec{D}_\beta = |\vec{C}_2 . \vec{X}_\beta - \vec{X}|,$$
$$\vec{D}_\delta = |\vec{C}_3 . \vec{X}_\delta - \vec{X}| \tag{5}$$

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1 . (\vec{D}_\alpha),$$
$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2 . (\vec{D}_\beta),$$
$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3 . (\vec{D}_\delta) \tag{6}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{7}$$

The iteration termination of the grey wolf optimization (GWO) calculation is based on the convergence of the optimal solution required. The best solution is obtained from calculating the alpha score, and the parameter that will be used in the next calculation or next step is the value of the alpha position that has been obtained. If the problem does not have any constraint tag then to get the solution by stopping the iteration process by setting the maximum iteration (t) at the beginning of the computation process [14].

2.3. Support vector machine (SVM)

Support vector machine (SVM) is a classification method introduced by Boser, Guyon, and Vapnik in 1992. The SVM method is extensively used in bioinformatics because of its high accuracy, and is able to handle data with high dimensions [15]. SVM aims to find a hyperplane between two target classes in the data and maximize the margin [16]. The linear model for hyperplane is in (8) [17].

$$f(x) = sign(w^T x + b) \tag{8}$$

Where,
w = weight vector;
b = bias term;
x = input vector.

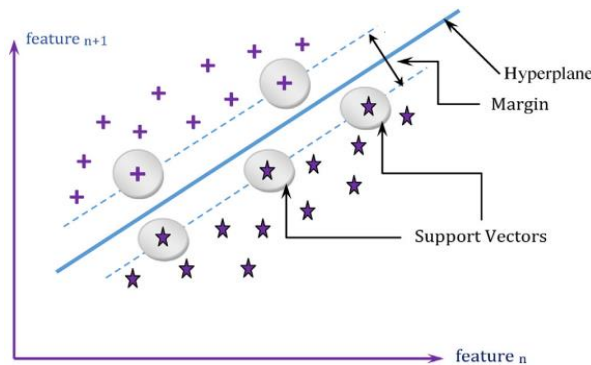The following Figure 1 is a representation of the SVM method [18].

Figure 1. A representation of SVM

Given a dataset $\{(x_i, y_i)| x_i \in X^n, y_i \in \{-1, +1\}, i = 1, 2, \ldots, D\}$, the SVM optimization problem is in (9) [17]:

$$\min. \frac{1}{2}||w||^2 \tag{9}$$

Subject to:

$$y_i[w_i^T x_i + b] \geq 1 \tag{10}$$

If the classification seems to have an error tolerance, the SVM optimization is based on (11) [17]:

$$\text{min.} \frac{1}{2}||w||^2 + C \sum_{i=1}^{D} \xi_i \tag{11}$$

Subject to:

$$\begin{cases} y_i[w_i^T x_i + b] \geq 1 - \xi_i \\ \xi_i \geq 0 \end{cases} \tag{12}$$

When the problem cannot be separated linearly in the input space, the strategy is to apply kernel function in SVM. The kernel function defined in (13).

$$K(x_i, x_j) = \langle \varphi(x_i), \varphi(x_j) \rangle \tag{13}$$

Where $\varphi(x_i)$ denotes the input data x_i's mapped feature space, and in the high-dimensional feature space, $K(x_i, x_j)$ is the kernel function that is equal to the inner product of two vectors [19]. Thus, the kernel SVM formula is in the (14) [20].

$$\min_{w, \xi} \frac{1}{2} w^T w + C \sum_{i=1}^{D} \xi_i \tag{14}$$

Subject to:

$$\begin{cases} y_i(w^T \varphi(x_i) + b) \geq 1 - \xi_i \\ \xi_i \geq 0, \forall i \end{cases} \tag{15}$$

We can get the SVM kernel optimization expressed in terms of variables λ_i [21]:

$$\max_{\lambda}(\sum_{i=1}^{D} \lambda_i - \frac{1}{2} \sum_{i=1}^{D} \lambda_i \lambda_j y_i y_j K(x_i, x_j) \tag{16}$$

Subject to:

$$\begin{cases} \sum_{i=1}^{D} \lambda_i y_i \\ 0 \leq \lambda_i \leq C \end{cases} \tag{17}$$

In the next algorithm, the kernel function that will be used is RBF kernel where

$$K(x_i, x_j) = \exp\left(-\gamma \|x_i - x_j\|^2\right) \tag{18}$$

2.4. Optimizing SVM parameters using GWO algorithm

SVM has been used successfully in solving many classification problems. SVM performance is affected by the choice of kernel function chosen and their parameter values [22]. As we know, the accuracy of prediction of the SVM model is related to the correct parameters. To maximize the SVM performance the parameters must be tuned; the error penalty C and the kernel parameters. To adjust the parameters (C, γ), the training data was used. Therefore, this meta-heuristic technique can be use to find the feasible SVM parameters as written on review from Faris, Aljarah, Al-Betar and Mirjalili [23]. Figure 2 depicts the process flow chart of the GWO-SVM algorithm [24], [25].

2.5. Performance measurement

To evaluate performance of the classification models, the score of accuracy will be used. Compute the accuracy score using the formula (19) with TP TN FN FP based on Table 2.

$$\text{Accuracy} = \frac{TN + TP}{FP + TP + TN + FN} \tag{19}$$

Where TP is number of tests are having acute sinusitis and accurately diagnosed having acute sinusitis and TN is number of tests are non-acute sinusitis and accurately diagnosed belonging non-acute sinusitis. While FN is number of tests with acute sinusitis but inaccurately classified as non-acute sinusitis and FP is number of tests are non-acute sinusitis but inaccurately classified as having acute sinusitis.

Table 2. Confusion matrix

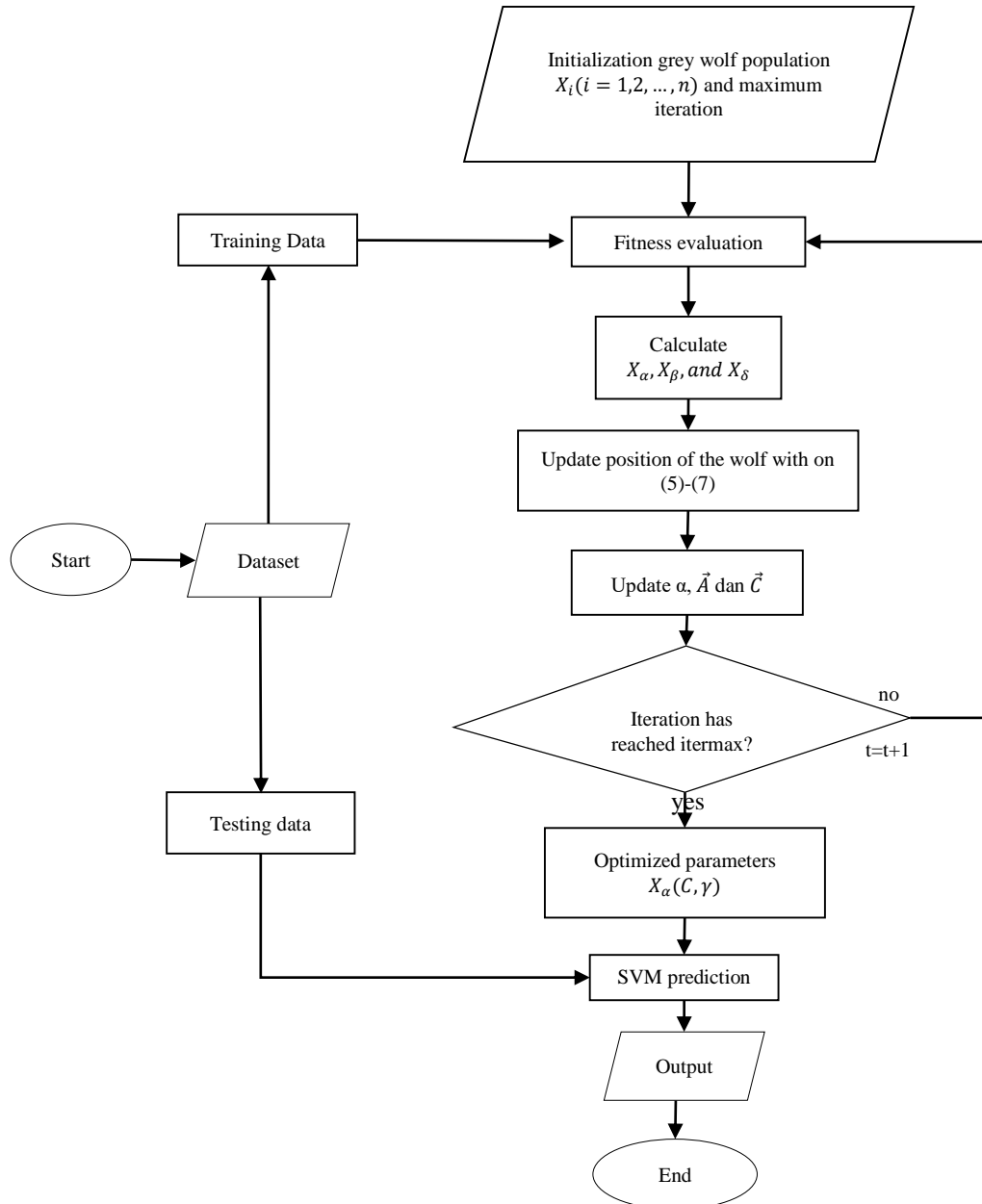|		Prediction	
		Acute Class	Non-acute Class
Actual	Acute Class	True Positive (TP)	False Negative (FN)
	Non-acute Class	False Positive (FP)	True Negative (TN)

Figure 2. Algorithm of GWO-SVM

3.   RESULTS AND DISCUSSION

The performance of SVM RBF Kernel and GWO-SVM RBF Kernel was analyzed using the python programming language with different percentages of training data, which started from 10% to 90%. The parameters obtained through GWO-SVM for each percentage of training data are shown in Table 3, and SVM RBF Kernel and GWO-SVM RBF Kernel results shown in Table 4 and Table 5. Figure 3 shows the comparative analysis for accuracy of SVM and GWO-SVM. The result showed that by optimizing the SVM classifier's input parameters, the proposed GWO-SVM solution outperformed the SVM algorithm.

Table 3. The parameters obtained through GWO-SVM With 20 search agents and 5 Iterations

Training Data (%)	C	γ
10	7.004445536344297	6.476843943302934
20	8.731737527898455	1.359474540632652
30	4.303597397541817	6.281462902997266
40	4.4446128650007	1.0827894101780806
50	3.4322191738424563	2.1017987982053046
60	2.0760205960583415	2.372396801652168
70	6.0559739832478385	2.6561410142050117
80	7.547317942283015	3.4721712260851993
90	6.899716798237926	2.5180009935420666

Table 4. The performance of SVM RBF kernel

Training Data (%)	Accuracy (%)
10	96.11
20	93.12
30	88.57
40	98.33
50	98.00
60	97.50
70	96.67
80	97.50
90	95.00

Table 5. The performance of GWO-SVM RBF kernel

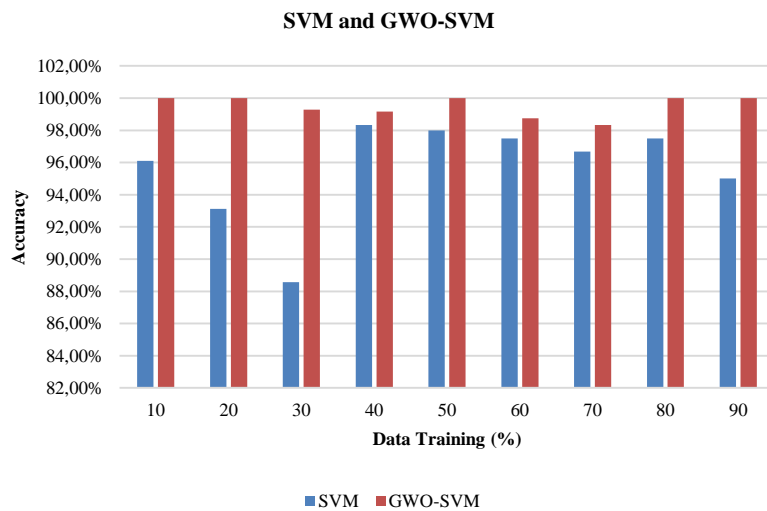| Training Data (%)	Accuracy (%)
10	100
20	100
30	99.28
40	99.17
50	100
60	98.75
70	98.33
80	100
90	100

Figure 3. Comparison between SVM RBF Kernel and GWO-SVM RBF Kernel

4. CONCLUSION

This research begins with acute sinusitis data, which modeled by proposed method GWO-SVM RBF Kernel and SVM RBF Kernel as a comparison, and the GWO technique used to optimizing the parameters in SVM. The proposed method produce the lowest accuracy of 98.33% when data training at 70% and achieved the highest accuracy of 100% when data training at 10%, 20%, 50%, 80% and 90%. The comparison result above illustrate the effectiveness of using GWO to tuned the parameters in SVM classifier. However, the difference in accuracy results are not very significant, because using SVM without GWO alone provides good accuracy results. For future research, this method can be used for other data with expectation that the accuracy of the classification results obtained is better than just using conventional methods, and also if it still use the SVM classifier, other kernel functions can be used.

REFERENCES

[1]   A. S. Arora and J. Singh, "A Framework for Enchancing the Thermographic Evaluation on Characteristic Areas for Pranasal Sinusitis Detection," In *Infrared Physics & Technology,* vol. 85, pp. 457-464, 2017, doi: 10.1016/j.infrared.2017.08.011.
[2]   B. Feldt, "Acute Sinusitis," in *Southern Medical Journal*, vol. 106, no. 10, pp. 577-581, 2013, doi: 10.1097/smj.0000000000000002.
[3]   R. M. Rosenfeld, "Acute Sinusitis in Adults," in *The New England Journal of Medicine*, 375:962-970, 2016. doi: 10.1056/NEJMcp1601749.
[4]   J. Arfiani et al,. "Kernel Spherical K-Means and Support Vector Machine for Acute Sinusitis Classification," in *IOP Conference Series Materials Science and Engineering,* 546 052011, 2019, doi: 10.1088/1757-899X/546/5/052011.
[5]   T. L. Octaviani et al., "Ovarian Cancer Classification using Bayesian Logic Regression," in *IOP Conference Series Materials Science and Engineering,* 546 052049, 2019. doi: 10.1088/1757-899X/546/5/052049.
[6]   Z. Rustam and S. Hartini "Classification of Breast Cancer using Fast Fuzzy Clustering based on Kernel," in in *IOP Conference Series Materials Science and Engineering,* 546 52067, 2019, doi: 10.1088/1757-899X/546/5/052067.
[7]   N. Salmi and Z. Rustam. "Naive Bayes Classifier for Predicting the Colon Cancer" in *IOP Conference Series Materials Science and Engineering,* 546 052068, 2019. doi: 10.1088/1757-899X/546/5/052068.
[8]   M. Huljanah et al., "Feature Selection using Random Forest Classifier for Predicting Prostate Cancer," in *IOP Conference Series Materials Science and Engineering,* 546 052031, 2019. doi: 10.1088/1757-899X/546/5/052031.
[9]   T. D. DenOtter and J. Schubert. "Hounsfiled Unit" in StatPearls Publishing, 2019.
[10]   S. Mirjalili et al., *"Grey wolf optimizer," Adv Eng Softw,* 69:46–61, 2014. doi: 10.1016/j.advengsoft.2013.12.007.
[11]   T. R. Biyanto et al., "Optimization of Acid Gas Sweetening Plant Based on Least Squares-Support Vector Machine (LS-SVM) Model and Grey Wolf Optimizer (GWO)" in *International Seminar on Sensors, Instrumentation, Measurement and Metrology (ISSIMM),* 2016, doi: 10.1109/ISSIMM2016.7803711.
[12]   Z. Mustaffa, M. H. Sulaiman and M. N. M. Kahar. "LS-SVM Hyper-parameters Optimization based on GWO Algorithm for Time Series Forecasting," in *International Conference on Software Engineering and Computer Systems (ICSECS),* 2015, doi: 10.1109/ICSECS.2015.7333107.
[13]   A. H. Sweidan et al., "Water Quality Classification Approach based on Bio-inspired Gray Wolf Optimization," in *7th International Conference of Soft Computing and Pattern Recognition (SoCPaR),* 2015, doi: 10.1109/SOCPAR.2015.7492777.
[14]   S. Eswaramoorthy, "Grey Wolf Optimization Based Parameter Selection for Support Vector Machines", in *The International Journal for Computation and Mathematics in Electrical and Electronic Engineering*, vol. 35, no. 5, pp. 1513-1523, 2016. doi: 10.1108/COMPEL-09-2015-0337.
[15]   B. Schölkopf, K. Tsuda and J. P. Vert. "Kernel Methods in Computational Biology," in *MIT Press series on Computational Molecular Biology,* 2004.
[16]   A. Benhur and J. Weston, "A user's guide to support vector machines," in *Methods in Molecular biology,* vol. 609, pp. 223-239, 2010, doi: 10.1007/978-1-60327-241-4_13.
[17]   N. Cristianini and J. S. Taylor, "An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods" in Cambridge Press University, 2000, doi: 10.1017/CBO9780511801389.
[18]   S. R. Naganna and P. C. Deka., "Support vector machine applications in the field of hydrology: A review," in *Applied Soft Computing,* vol. 19, pp. 372-386, 2014, doi: 10.1016/j.asoc.2014.02.002.
[19]   X.-Q. Bian et al., *"Prediction of Sulfur Solubility in Supercritical Sour Gases Using Grey Wolf Optimizer-Based Support Vector Machine," in Journal of Molecular Liquids*, vol. 261, pp. 431–438, 2018, doi: 10.1016/j.molliq.2018.04.070.
[20]   M. Awad and R. Khanna. "Efficient Learning Machine," in *Springer Science + Business Media New York*, p 39-55, 2015, doi: 10.1007/978-1-4302-5990-9.
[21]   J. Xiao and R. Zhuo, "Research on Motor Rolling Bearing Fault Classification Method Based on CEEMDAN and GWO-SVM," *2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, 2018, pp. 1123-1129, doi: 10.1109/IMCEC.2018.8469276.
[22]   A. L. D. Rossi and A. C. P. L. F. Carvalho. "Bio-inspired Optimization Techniques for SVM Parameter Tuning" in *10th Brazilian Symposium on Neural Networks,* 2008, doi: 10.1109/SBRN.2008.28.
[23]   H. Faris, *et al.* "Grey Wolf Optimizer: A Review of Recent Variants and Applications" in *Neural Computing and Applications*, 2019, doi: 10.1007/s00521-017-3272-5.
[24]   X.-Q. Bian, et al., "A grey wolf optimizer-based support vector machine for the solubility of aromatic compounds in supercritical carbon dioxide," in *Chemical Engineering Research And Design*, vol. 123, pp. 284–294, 2017, doi: 10.1016/j.cherd.2017.05.008.
[25]   S. Jiang, et al., "Hybrid Genetic Grey Wolf Algorithm for Large-Scale Global Optimization," in *Compexity,* vol. 2019, 2019, Art. no. 2653512, doi: 10.1155/2019/2653512.

BIOGRAPHIES OF AUTHORS

Ajeng Maharani Putri is a senior student at the University of Indonesia's Department of Mathematics. Ms. Ajeng is passionately studying data mining and machine learning.

Zuherman Rustam is an Associate Professor and a lecturer in intelligence computation at the University of Indonesia's Department of Mathematics. He received his Master of Science in informatics from Paris Diderot University in 1989, and his Ph.D. in computer science from the University of Indonesia in 2006. Assoc. Prof. Dr. Rustam is an IEEE member who is actively researching artificial intelligence, neural network, machine learning, and pattern recognition.

Jacub Pandelaki is an academic senate and lecturer at the Faculty of Medicine, University of Indonesia. He earned his medical doctorate from the Faculty of Medicine, University of Indonesia in 1989 and his Ph.D. from the same university in 2010. Dr. dr. Jacub Pandelaki, Sp.Rad is a doctor and the chair of the International Radiology Division at the Department of Radiology at Cipto Magunkusumo Hospital in Indonesia.

Ilsya Wirasati is a final year student at the University of Indonesia's Department of Mathematics, who is currently working on her thesis. Her research is firmly about applied mathematics using machine learning in medical field. Ms. Ilsya's specialties in research are mostly about machine learning, mathematical modeling, and data mining.

Sri Hartini is a Bachelor of Science from the Department of Mathematics, University of Indonesia. Ms. Hartini is a machine learning, computer vision, neural networks, and deep learning researcher who works in a number of fields.