Comparison some of kernel functions with support vector machines classifier for thalassemia dataset

Ilsya Wirasati, Zuherman Rustam, Jane Eva Aurelia, Sri Hartini, Glori Stephani Saragih
Department of Mathematics, University of Indonesia, Indonesia

Article Info

ABSTRACT

In the medical field, accurate classification of medical data is really important because of its impact on disease detection and patient's treatment. Technology, machine learning, is needed to help medical staff to improve accuracy to classify disease. This research discussed some kernel functions, such as gaussian radial basis function (RBF) kernel, Polynomial kernel, and linear kernel with support vector machine (SVM) to classify thalassemia data. Thalassemia is a genetic blood disorder which is also one of the major public health problems. In this paper, there is an explanation about thalassemia, SVM, and some of the kernel functions that serve as a comprehensive source for the next research about this topic. Furthermore, there is a comparison result from three kernel functions to find out which one has the best performance. The result is gaussian RBF kernel with SVM is the best method with an average of accuracy 99,63%.

Corresponding Author:

Ilsya Wirasati
Department of Mathematics
University of Indonesia
Jl. Prof. DR. Sudjono D. Pusponegoro, Pondok Cina, Depok, Jawa Barat 16424, Indonesia
Email: ilsya.wirasati@ui.ac.id

1. INTRODUCTION

Thalassemia is one of the main public health problems with highly prevalent in the area extending from sub-Saharan Africa, through the Mediterranean region and Middle East, to the Indian subcontinent and East and Southeast Asia [1], [2]. However, migrations of people caused thalassemia genes to spread throughout the world and extend to Indonesia. There are 7% of the world's population as carriers of thalassemia with the death of about 50,000-100,000 children [3]. In Indonesia, thalassemia is one of the most common chronic diseases [4]. Currently, thalassemia ranks 5th among non-communicable diseases after heart disease, cancer, kidney, and stroke with the number of carriers 3.8% of the entire population in Indonesia. Based on data from the Indonesian Thalassemia Foundation, there has been a steady increase in thalassemia cases from 2012 until 2018 [3].

Thalassemia is a genetic disease because of blood disorders inherited from family. Thalassemia sufferers' body makes an abnormal form or an inadequate amount of hemoglobin [1], [5]. Hemoglobin allows red blood cells to carry oxygen [6]. When there is not enough hemoglobin, the body's red blood cells do not function properly, and they die more quickly. And then, the oxygen delivered to all the other cells of the body is not enough.

The cause of thalassemia is mutations in the DNA of cells that make hemoglobin [7]. Hemoglobin is made of two different parts, called alpha and beta. Therefore, there are two types of thalassemia, such as alpha-thalassemia or beta-thalassemia. According to [8], the new classification has been simplified based on the way of treatment namely non-transfusion-dependent thalassemia (NTDT) and transfusion-dependent

thalassemia (TDT). Because of differences in treatment, early detected thalassemia with a screening process is necessary to help thalassemia suffers to get the right treatment. The aim is to increase their life expectancy and reduce the risk of thalassemia to the next generation. Thus, it is important to obtain a precise thalassemia diagnosis.

Nowadays, in healthcare, it is significant to invest the development in computer technology to enhance processing the medical data [5]. Machine learning technologies, one of computer technology, can help us in classification problems on large datasets. It has an important role because it can be applied in daily life such as biomedical data. However, there are several interesting challenges recently such as our data may come from multiple heterogeneous sources, our data may have a huge number of samples and require a method to understand the complex model, and our data may have few samples but lie in high dimension and is spatiotemporal. New developments in statistics and kernel methods is required to these challenges [9].

There are some methods on previous researches to classify thalassemia, such as fuzzy kernel robust C-means, fuzzy C-means, and fuzzy kernel C-means [4], neural networks and genetic programming [10], artificial intelligence algorithms [11], artificial neural network [12], and naïve bayes [13]. Also, [12], [14] used SVM that showed good result with 93.2% accuracy and 100% AUC respectively.

This research used some of kernel functions with support vector machine (SVM) to classify thalassemia. SVM can be modified with various kernel functions, as an essential component, to get a better result. Therefore, a comparison between that essential component for classifying thalassemia should be done. It will help the medical staff to overcome the classification problems. This research discussed some of the kernel functions such as the linear kernel, polynomial kernel, and gaussian radial basis kernel. The aim is to find out which kernel function that gives the highest accuracy for classifying thalassemia in the SVM method.

2. RESEARCH METHOD

Support vector machines (SVM) is supervised machine learning. Originally, SVM algorithm proposed by Vapnik and Lerner [15], [16]. SVM can be applied for classification and regression [17], [18]. It claimed that SVM is a method that has a high accuracy for classification [19]. Mapping form input space to a higher dimensional space is the idea of SVM. SVM constructs a hyperplane to separate data into classes [20]. The selected hyperplanes are those that maximize the margin of classification edges [21].

Let $\{x_i, y_i\}_i^N$ is the dataset where, $x_i \epsilon R^D$ is feature of vector, y_i is class label for x_i, and N is the number of samples. To find the best hyperplane, this is main formula of support vector machines:

$$f(x) = \boldsymbol{w} \cdot \boldsymbol{x} + b \tag{1}$$

That formula contains **w** (weight) as the orthogonal vector to the hyperplane determining its orientation, b (bias) as the distance from the origin to the hyperplan, and **x** indicates the training sample [22]. The aim is to maximize the margin.

Moreover, SVM goal is construct the two planes, let say H1 and H2, as (2) and (3):

$$H_1 \longrightarrow \boldsymbol{w}^T \boldsymbol{x}_i + b = +1 \, for \, \boldsymbol{y}_i = +1 \tag{2}$$

$$H_2 \longrightarrow \boldsymbol{w}^T \boldsymbol{x}_i + b = -1 \, for \, \boldsymbol{y}_i = -1 \tag{3}$$

where the plane for the positive class is $\boldsymbol{w}^T \boldsymbol{x}_i + b \geq +1$ is and the plane for the negative class is $\boldsymbol{w}^T \boldsymbol{x}_i + b \leq -1$. See Figure 1 illustrate the hyperplane in SVM. The problem of SVM optimization can be written as:

$$\text{Minimize } \frac{1}{2}\|w\|^2 \tag{4}$$

$$\text{s.t. } y_i \left(\boldsymbol{w}^T \cdot x_i + b \right) \geq 1, \forall i = 1, \dots, N \tag{5}$$

By solving the problem above, formula of $\boldsymbol{w}$ and b can be written as:

$$\boldsymbol{w} = \sum_{i=1}^N a_i y_i \boldsymbol{x_i} \tag{6}$$

$$b = \frac{1}{N_s} \sum_{i \epsilon S} (y_i - \sum_{m \epsilon S} a_m y_m \boldsymbol{x}_m) \tag{7}$$

Then, decision formulas of SVM can be written as:

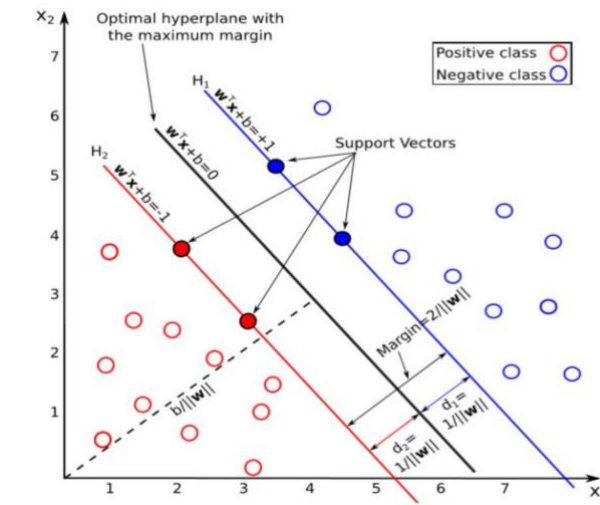$$f(x) = sign(\boldsymbol{w} \cdot \boldsymbol{x} + b) \tag{8}$$

Figure 1. Illustration of SVM [16]

SVM has several advantages, such as its capability to process data with large amounts in high dimensions [23]. Also, SVM implemented easily using linear boundaries as shown in Figure 1. However, there are classification problems where can not using a linear boundary to separate classes [24]. See Figure 2, that case is non-linear separable data. The best way to approach a non-linear decision boundary is to expand the original feature space. Nevertheless, it makes computations intractable because the original feature is enlarged to high dimensional space. To tackle that issue, we applied the 'kernel trick' using a kernel function.
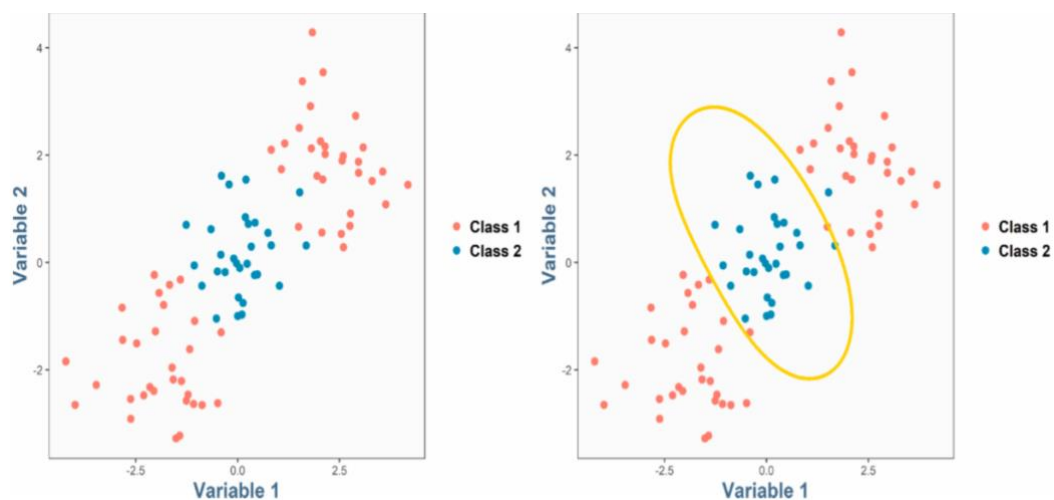

Figure 2. Non-linear separable data [25]

SVM classification performance closely relies on the kernel function [26]. Therefore, a kernel function is the most essential component to make the SVM method get higher accuracy [27]. When a task is difficult in the original problem space, kernel function helps to transform input space into another space where we can work easier [25]. On another word, kernel function work for transforming data into a higher-dimensional space [28], [29]. Its approach is mapping data into kernel space where data become linearly separable [26].

The kernel function can be written as:

$$\kappa(x_i, x_j) =< \varphi(x_i), \varphi(x_j) > \tag{9}$$

Example, we construct *lifting map* $\varphi: \chi \to \mathcal{H}$ with $\varphi: (x_1, x_2) \to (x_1^2 + \sqrt{2}x_1x_2 + x_2^2)$. This map *lifting* the data from $\chi = \mathbb{R}^2$ to $\mathcal{H} = \mathbb{R}^3$ [30]. Therefore, φ mapping data from dimensional space to feature space.

The problem of SVM optimization will be as follows:

$$\text{Minimize} \frac{1}{2}\|w\|^2 + C \sum_{i=1}^N \epsilon_i$$

$$\text{s.t} y_i \left(\boldsymbol{w}^T \cdot \varphi(\boldsymbol{x_i}) + b\right) - 1 + \epsilon_i \geq 0, \forall i = 1, \dots, N$$

By solving the problem above, formula of $\boldsymbol{w}$ and b will be as (10) and (11):

$$\boldsymbol{w}^* = \sum_{i=1}^N a_i y_i \varphi(\boldsymbol{x_i}) \tag{10}$$

$$b^* = \frac{1}{N_s}\sum_{i\epsilon S}(y_i - \sum_{m\epsilon S} a_m y_m \varphi(\boldsymbol{x_m})) \tag{11}$$

Then, decision formulas of SVM will be as (12):

$$f(x) = sign(\boldsymbol{w}^* \cdot \varphi(\boldsymbol{x_i}) + b^*) \tag{12}$$

where ϵ_i is slack variable or measure of the misclassification errors which should be minimize. C is the penalty or determines the trade-off between the minimization of error and the maximization of the classification margin.

In this research, authors proposed three kernels which applied for thalassemia classification:
a. Gaussian radial basis kernel

$$\kappa(x_i, x_j) = \exp\frac{\|x_i - x_j\|^2}{2\sigma^2} \tag{13}$$

Wheres σ is the only parameter that defines width kernel. Its impact to close or far a single training sample reaches. Also, σ can defined as the radius of influence of samples which is affected by the classification model. From research in [16], a small σ indicates the width of the kernel is small so the model focuses on a small set of data and the new hypersurface will be spiky. It may leads to an overfitting problem. The opposite, a high σ increases the kernel width and then most of the data are transformed into a flat hyperspace which leads to the underfitting problem.
b. Polynomial kernel

$$\kappa(x_i, x_j) = (< x_i, x_j > +1)^d \tag{14}$$

Wheres d is degree of polynomial kernel function. From research in [16], high degree would increase the complexity of the classification model. It can be seen as overfitting problem because testing error increases but training error decreases. The opposite, with a small d may leads to a high bias and low variance or underfitting problem.
c. Linear kernel:

$$\kappa(x_i, x_j) = x_i^T x_j \tag{15}$$

This kernel function is the simplest kernel function which the results of learning algorithms are often equivalent to SVM without kernel functions [16]. By comparing these kernels, the expectation is we know which kernel gives the highest accuracy. To calculate the accuracy, a confusion matrix is used. The formula for accuracy is:

$$accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{16}$$

$$precision = \frac{T_P}{T_P + F_P} \tag{17}$$

$$recall = \frac{T_P}{T_P + F_N} \tag{18}$$

$$F1\ Score = \frac{2\ x\ precision\ x\ recall}{precision + recall} \tag{19}$$

T_P: Number of samples having thalassemia classified correctly.
F_P: Number of healthy people that were incorrectly classified to thalassemia.
F_N: Number of samples with thalassemia that were incorrectly classified as healthy.
T_N: Number of healthy individuals correctly spotted.

3. RESULTS AND DISCUSSION

In this paper, thalassemia data received from Harapan Kita Children and Women's Hospital, Indonesia, and it consist of 150 samples. The dataset of thalassemia represented by 10 variables such as Hemoglobin (g/dL), Haematocrit Percent (%), Leukocyte Count ($103/\mu L$), Basophils Percent (%), Eosinophils Percent (%), Rod Neutrophils Percent (%), Segment Neutrophils Percent (%), Lymphocytes Percent (%), Monocytes Percent (%), and Platelet Counts ($103/\mu L$). By default authors utilized the Shapiro-Wilk algorithm to assess the normality of the distribution of instances with respect to the feature. A barplot as shown in Figure 3, is then drawn showing the relative ranks of each feature. Platet Counts has the highest ranking.
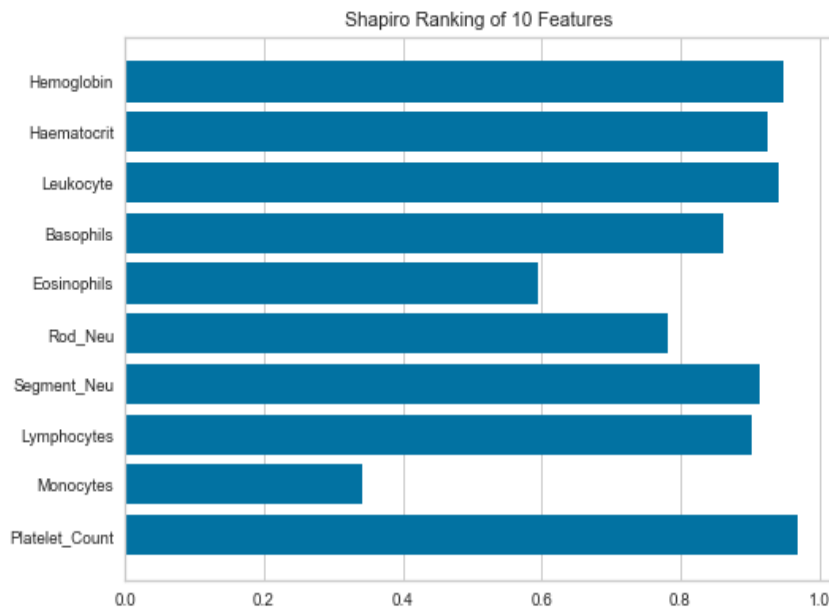

Figure 3. Ranking of thalassemia data features with saphiro algorithm

This research used training data diverse from 10% to 90% and used $\sigma = 0.1$ for Gaussian RBF kernel and d=3 for polynomial kernel. The reason is, from the number of the experiment that is obtained, $\sigma = 0.1$ and d=3 has the best performance. This chosen $\sigma = 0.1$ is also supported by [16].

It is shown in Table 1, the SVM model with a gaussian radial basis function kernel produces the best accuracy for classifying thalassemia data with an average of accuracy 99.63%. The second-best is a linear kernel with 98.23% accuracy. The last one is a polynomial kernel with 97.9% accuracy. Linear kernel SVM has the best accuracy of 100% with 10% and 30% training data. On the other side, the polynomial kernel has the best accuracy of 100% if the model uses 10-30% and 50% training data. And for gaussian radial basis function gives the best accuracy with 10-50%, 70%, and 80% training data. For F1 Score, gaussian radial basis still the best one. In Table 2, the gaussian radial basis kernel gives the best performance with an average precision of 99.56% and an average recall of 99.78%. However, there is a difference in second place between

precision and recall. SVM linear is in second place for precision, while for recall SVM polynomial is in second place.

Table 1. The accuracy and F1 score of SVM with kernel function

Training Data	Accuracy			F1 Score		
	SVM Linear	SVM Polynomial	SVM Gaussian	SVM Linear	SVM Polynomial	SVM Gaussian
10%	100.00	100.00	100.00	100.00	100.00	100.00
20%	96.67	100.00	100.00	96.00	100.00	100.00
30%	100.00	100.00	100.00	100.00	100.00	100.00
40%	98.00	96.67	100.00	98.00	96.00	100.00
50%	98.67	100.00	100.00	99.00	100.00	100.00
60%	98.89	97.77	98.89	99.00	98.00	99.00
70%	97.14	96.19	100.00	98.00	96.00	100.00
80%	99.16	97.5	100.00	98.00	100.00	100.00
90%	95.56	93.33	97.78	97.00	93.00	99.00
Average	98.23	97.94	99.63	98.33	98.11	99.78

Table 2. The precision and recall of SVM with kernel function

Training Data	Precision			Recall		
	SVM Linear	SVM Polynomial	SVM Gaussian	SVM Linear	SVM Polynomial	SVM Gaussian
10%	100.00	100.00	100.00	100.00	100.00	100.00
20%	100.00	100.00	100.00	93.00	100.00	100.00
30%	100.00	100.00	100.00	100.00	100.00	100.00
40%	100.00	100.00	100.00	95.00	94.00	100.00
50%	98.00	100.00	100.00	98.00	100.00	100.00
60%	100.00	100.00	100.00	98.00	96.00	98.00
70%	95.00	95.00	100.00	98.00	98.00	100.00
80%	100.00	100.00	100.00	94.00	96.00	100.00
90%	100.00	97.00	96.00	95.00	92.00	100.00
Average	99.22	99.11	99.56	96.78	97.33	99.78

Authors also used other machine learning, such as KNN with k=7 and random forest. The result is 90% accuracy from KNN and 100% accuracy from random forest. Nevertheless, SVM with some of the kernel functions still give the highest accuracy, 100%, so it can be said that SVM performed the best machine learning method to classify thalassemia.

4. CONCLUSION

Machine learning can help medical staff to classify thalassemia disease precisely. If early detection is done, patients can get the right treatment. It helps them increase their life expectancy and reduce the risk of thalassemia to the next generation. In this research, there are three kernel functions used in SVM with linear, polynomial, and gaussian radial basis function kernel. Kernel function can help SVM to transform input space into a higher-dimensional space where we can work easier.

From this research, support vector machine with gaussian RBF kernel is the best one to classify thalassemia data from Harapan Kita Children and Women's Hospital, Indonesia. We can see in Table 1, each kernel performs the highest accuracy. However, if we see the average accuracy, gaussian RBF is the best one with an accuracy of 99.63%. The second-best is a linear kernel with 98.23% accuracy. The last one is a polynomial kernel with 97.9% accuracy. Besides that, the gaussian radial basis also gives the highest F1 score of 99.78%. Also, in Table 2, the gaussian RBF kernel has the highest average of precision and recall with 99.56% and 99.78% respectively. For future research, use a larger dataset is recommended to generate higher accuracies in each method. Also, we believe that future research can develop this method to give the best accuracy for predicting or classifying other diseases.

REFERENCES

[1]   R. Risoluti *et al*., "Update on thalassemia diagnosis: New insights and methods," in *Talanta*, vol. 183, pp. 216-222, 2018, doi: 10.1016/j.talanta.2018.02.071.

[2]   A. T Taher, D. J. Weatherall, and M. D. Cappellini, "Thalassemia," in *The Lancet*, vol. 391, pp.157-167, 2018, doi: 10.1016/S0140-6736(17)31822-6.

[3]   "Ministry of health republic of Indonesia," http://www.kemkes.go.id/ [accessed 2020 Feb 11].

[4]   Z. Rustam *et al*., "Comparison of Fuzzy C-Means, Fuzzy Kernel C-Means, and Fuzzy Kernel Robust C-Means to Classify Thalassemia Data," in *International Journal on Advanced Science, Engineering and Information Technology*, vol. 9, no. 4, pp. 1205-1210, 2019, doi: 10.18517/ijaseit.9.4.9580.

[5]   N. C. Egejuru *et al*., "Using Data Mining Algorithms for Thalassemia Risk Prediction," in *International Journal of Biomedical Science and Engineering,* vol. 7, no. 2, pp. 33-44, 2019, doi: 10.11648/j.ijbse.20190702.12.

[6]   S. Fucharoen and P. Winichagoon, "Thalassemia and abnormal hemoglobin," in *International journal of hematology* vol. 76, no. 2, pp. 83-89, 2002, doi: 10.1007/BF03165094.

[7]   H. Nawal *et al*., "Thalassemia review: features, dental considerations and management," in *Electronic physician*, vol. 9, no. 3, pp. 4003-4005, 2017, doi: 10.19082/4003.

[8]   V. Viprakasit and S. Ekwattanakit, "Clinical classification, screening and diagnosis for thalassemia," in *Hematology/Oncology Clinics*, vol. 32, no. 2 pp. 193-211, 2018. doi: 10.1016/j.hoc.2017.11.006

[9]   H. H. Zhou, "Kernel Methods in the Analysis of Big and Complex Data: A Modern Statistical Challenge," Ph.D dissertation, Wisconsin-Madison Univ., Madison, United State, 2019. [Online]. Available: https://search.proquest.com/openview/0b85eaa4c7718832bf36bb357ec7f1ac/1?pq-origsite=gscholar&cbl=18750&diss=y

[10]   Wongseree *et al*., "Thalassemia classification by neural networks and genetic programming," in *Information Sciences*, vol. 177, no. 3, pp. 771-786, 2007, doi: 10.1016/j.ins.2006.07.009.

[11]   F. Yousefian, T. Banirostam, and A. Azarkeivan, "Prediction Thalassemia Based on Artificial Intelligence Techniques: A Survey," in *International Journal of Advanced Research in Computer and Communication Engineering*, vol. 6, no. 6, pp. 281-287, 2017, doi: 10.17148/IJARCCE.2017.6847.

[12]   S. R. Amendolia *et al*., "A Real-Time Classification System of Thalassemic Pathologies Based on Artificial Neural Networks," in *Medical Decision Making*, vol. 22, no. 1, pp. 18-26, 2002, doi: 10.1177/0272989x0202200102.

[13]   D. Setsirichok *et al*., "Classification of complete blood count and haemoglobin typing data by a C4.5 decision tree, a Naïve Bayes classifier and a Multilayer Perceptron for Thalassaemia screening," in *Biomedical Signal Processing and Control*, vol. 7, no. 2, pp. 202-212, 2012, doi: 10.1016/j.bspc.2011.03.007.

[14]   V. Laengsri *et al*., "ThalPred: a web-based prediction tool for discriminating thalassemia trait and iron deficiency anemia," *BMC medical informatics and decision making*, vol. 19, 2019, Art. no 212, doi: 10.1186/s12911-019-0929-2.

[15]   A. Muzaffer *et al*., "Kernel Selection for Support Vector Machines for System Identification of a CNC Machining Center," in *IFAC-PapersOnLine*, vol. 52, no. 29, pp. 192-198, 2019, doi: 10.1016/j.ifacol.2019.12.643.

[16]   A. Tharwat, "Parameter investigation of support vector machine classifier with kernel functions," in *Knowledge and Information Systems,* vol. 61, pp. 1269-1302, 2019, doi: 10.1007/s10115-019-01335-4.

[17]   X. Zhang, J. Xiao and F. Gu, "Applying Support Vector Machine to Electronic Health Records for Cancer Classification," in *Spring Simulation Conference (SpringSim)*, pp. 1-9, 2019, doi: 10.23919/SpringSim.2019.8732906.

[18]   H. Zheng *et al.,* "Online Support Vector Machine with Adaptive Kernel Functions," in *International Conference on Computer Information Analytics and Intelligent Systems*, 2019, doi: 10.25236/ciais.2019.029.

[19]   Z. Rustam and N. P. N. A. Ariantari, "Support Vector Machines for Classifying Policyholders Satisfactorily in Automobile Insurance," in *Journal of Physics: Conference Series 1028 012005*, 2018, doi: 10.1088/1742-6596/1028/1/012005.

[20]   Z. Rustam and D. Zahras, "Comparison between Support Vector Machine and Fuzzy C-Means as Classifier for Intrusion Detection System," in *J. Phys.: Conf. Ser. 1028 012227*, 2018, doi: 10.1088/1742-6596/1028/1/012227.

[21]   M. Vatankhah, V. Asadpour and R. Fazel-Rezai, "Perceptual pain classification using ANFIS adapted RBF kernel support vector machine for therapeutic usage," in *Applied Soft Computing*, vol. 13, no. 5, pp. 2537-2546, 2013, doi: 10.1016/j.asoc.2012.11.032.

[22]   J. Abukhait and M. Obeidat, "Classification based on Gaussian-kernel Support Vector Machine with Adaptive Fuzzy Inference System," in *Przegląd Elektrotechniczny*, vol. 1, no. 5 pp. 16-24, 2018, doi: 10.15199/48.2018.05.03.

[23]   A R. Bagasta *et al.,* "Comparison of Cubic SVM with Gaussian SVM: Classification of Infarction for detecting Ischemic Stroke," in *IOP Conf. Series: Materials Science and Engineering 546 052016*, 2019, doi: 10.1088/1757-899X/546/5/052016

[24]   M. A Ohemeng, "A Comparative Study on Support Vector Machines," Ph.D dissertation, Mathematics, Nevada Univ., Reno, 2017. [Online]. Available: https://scholarworks.unr.edu/handle/11714/2816

[25] A. Apsemidis, S. Psarakis, and J. M. Moguerza. "A review of machine learning kernel methods in statistical process monitoring," in *Computers & Industrial Engineering*, vol. 142, 2020, Art. no. 106376, doi: 10.1016/j.cie.2020.106376.

[26] D. Zhao *et al.*, "Whale optimized mixed kernel function of support vector machine for colorectal cancer diagnosis," in *Journal of Biomedical Informatics*, vol. 92, 2019, Art. no. 103124, doi: 10.1016/j.jbi.2019.103124.

[27] Z. Liu and H. Xu, "Kernel Parameter Selection for Support Vector Machine Classification," in *Journal of Algorithms & Computational Technology*, vol. 8, no. 2, pp. 163-177, 2013, doi: 10.1260/1748-3018.8.2.163.

[28] T. V. Rampisela and Z Rustam, "Classification of Schizophrenia Data Using Support Vector Machine (SVM)," in *J. Phys.: Conf. Ser. 1108 012044*, 2018, doi: 10.1088/1742-6596/1108/1/012044.

[29] S. Hartini and Z. Rustam, "Hierarchical Clustering Algorithm Based on Density Peaks using Kernel Function for Thalassemia Classification," in *J. Phys.: Conf. Ser. 1417 012016*, 2019, doi: 10.1088/1742-6596/1417/1/012016.

[30] J. Moeller, "Kernels and Geometry of Machine Learning," Ph.D dissertation, Utah Univ., Utah, 2017. [Online]. Available: https://arxiv.org/abs/2010.15110

BIOGRAPHIES OF AUTHORS

Ilsya Wirasati is a final year student in the Department of Mathematics, University of Indonesia, 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.

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

Jane Eva Aurelia was born in Jakarta, 19 June 1998. She is a final year student in the Department of Mathematics, University of Indonesia. She is currently working on her thesis, which is firmly about applied mathematics using machine learning. Also, Ms. Jane'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, who is also completing the Master of Science at the University of Indonesia and is currently pursuing a Ph.D. in intelligence computation. Ms. Hartini is passionately researching machine learning, computer vision, neural networks and deep learning in various fields.

Glori Stephani Saragih was born in Medan, 17 January 1997. She is a Bachelor of Science from Department of Mathematics, Universitas Indonesia, who is completing the Master of Science at Universitas Indonesia and is currently pursuing a Ph.D. in intelligence computation. Ms. Glori is currently a Process Improvement Manager in PT. Aplikasi Karya Anak Bangsa (Gojek). Her current research is machine on machine learning and neural network in various fields, especially medical and finance.