Svm Performance Optimization Using PSO for Breast Cancer Classification

Rizki Habibi

Akademi Informatika dan Komputer Medicom, Indonesia
pakhabibi@gmail.com

Abstract: *Breast cancer is a very serious disease and requires the sufferer to undergo an intensive examination. This study proposes an optimization method for SVM parameters using the particle swarm optimization algorithm to classify breast cancer. The test results show that the classification using SVM with optimization using PSO is able to improve accuracy better than the classification using SVM without optimization, namely the determination of the parameters randomly. Breast cancer data classification accuracy increased to 78.91%. The best parameter values for c, γ, r, and d are 0.252101, 0.053248, 1 and 5, respectively. In this study, the polynomial and RBF kernels in general were able to produce higher accuracy than linear and sigmoid kernels.*
Keywords: *PSO; breast cancer; sym performance*

I. Introduction

Breast cancer is a cell disorder characterized by uncontrolled growth of cells in the breast. These cancer cells are most often suffered by women in both developed and developing countries. According to the Ministry of Health, the highest mortality rate for women is caused by breast cancer, which is 42.1 per 100,000 populations with an average death rate of 17 per 100,000 female populations (Ministry of Health, 2019). In 2018, WHO estimated that worldwide more than 627,000 women died from breast cancer (WHO, 2019). There are two types of breast cancer malignancy, namely malignant and belign. In general, the detection of the level of malignancy of breast cancer is by means of prognosis. The prognosis is the medical team's "best guess" in determining whether or not a patient is cured of breast cancer (Rachman and Purnami, 2012). Apart from prognosis, another way is the use of bioinformatics by using data mining techniques (Salama et al, 2012), because it has been proven to be able to detect the level of malignancy of breast cancer.

Currently, various data mining methods and algorithms that are applied for the development of data classifiers are known. The most well-known data mining tools, such as neural networks, decision trees, Support Vector Machine (SVM) algorithms, kNN (k Nearest Neighbors) algorithms, Parzen window algorithm, and others can be used in solving data classification problems. SVM is a machine learning method that works on the principle of Structural Risk Minimization (SRM) with the aim of finding the best hyperplane that separates two classes in the input space (Santoso, 2007). In addition, SVM also aims to minimize the upper limit of general errors. Another advantage of using SVM is that this method can be analyzed theoretically using the concept of computational learning theory.

In a study (Demidova et al., 2017) a hybrid and modified version of the PSO algorithm was applied to improve the search characteristics of the classic PSO algorithm in the development problem of the SVM classifier has been offered and investigated. With this the two hybrid versions of the PSO algorithm assume the use of the classic "Grid Search" (GS) algorithm and the corresponding "Design of Experiment" (DOE) algorithm, and a modified version of the PSO algorithm realizes the simultaneous search of the kernel function types, the parameter values of the kernel functions., and also regularization

28

parameter values. In addition, the question of the application of the Nearest Neighbor (kNN) algorithm in the development of the SVM classifier has been considered.

From this information, the importance of early detection and diagnosis of breast cancer has led many researchers to carry out research in the diagnosis of breast cancer and base the authors to conduct research on the Data Mining Algorithm in the diagnosis of breast cancer using the SVM and PSO methods using data from computationally calculated breast biopsy in under the microscope. The annual incidence of cervical cancer in Indonesia is estimated at 20,998 cases, and the number of deaths from cervical cancer is 9,498 people (ICO in Arysha, 2020). The data used in this study is data taken from data from Wicoxsin University UCI Machine Learning breast cancer patients. This paper is written in 4 sections. The second section describes the classification method used in this study, namely SVM and PSO as the optimal SVM parameter finding algorithm. The third section discusses the experimental results and the fourth section concludes the contents of the paper and suggestions for future research.

II. Research Methods

2.1 Support Vector Machine

Support Vector Machine (SVM) is an algorithm that works with nonlinear mapping which functions to transform initial training data into new, higher dimensions. In this new dimension, SVM will find the optimum linear hyperplane. By doing nonlinear mapping to higher dimensions, data from two classes will always be separated by a hyperplane. This method will find the hyperplane using support vectors and margins (Han et al., 2011). This method was first presented in 1992 by Vapnik, Boser, and Guyon at a Workshop on Computational Learning. SVM theory introduces a new strategy by finding the best hyperplane in the input space. The principle of SVM was originally a linear classifier, but SVM was later developed to be able to work on non-linear problems by including the kernel. The development of this SVM stimulates research interest in the field of pattern recognition in developing the potential capabilities of the SVM method both from a theoretical and application perspective. Recently, SVM has been successfully applied in solving practical problems. The concept of SVM can be explained simply as an effort to find the best hyperplane which functions as a separator of two classes in the input space.

Classification problems can be interpreted as trying to find the line that separates the two groups, as Figure 1 shows. The best separator hyperplane between the two classes can be found by measuring the hyperplane margin and finding its maximum point. Margin is the distance between the hyperplane and the closest pattern from each class. This closest pattern is called a support vector. The effort to find the location of this hyperplane is at the core of the SVM learning process. In what is assumed it is given a set S of points $x_i \in \mathbb{R}^n$ IRn where i = 1, 2, ..., N. Each point x_i belongs to one of the two classes and is thus labeled $Y_i \in \{1, -1\}$. Its purpose is to define a hyperplane equation that divides S leaving all points of the same class on the same side while maximizing the minimum distance between one of the two classes and the hyperplane. For this purpose some preliminary definitions are required (Pontil & Verri 1997).

First, the set S can be separated linearly if there is $w \in \mathbb{R}^n$ and $b \in \mathbb{R}$ such that:

$$x_i.w + b \geq +1 \text{ if } y_i = +1 \tag{1}$$

$$x_i.w + b \leq -1 \text{ if } y_i = -1 \tag{2}$$

w　　= weight vector perpendicular to the hyperplane (normal plane)
b　　= the position of the plane relative to the center of the coordinates
In simpler notation, the two inequalities above can be rewritten:

$$Y_i(w.x_i + b) \geq 1, \tag{3}$$

For i=1,2,…., N. The pair (w, b) represents the hyperplane of the following equation:

$$w.x + b = 0 \tag{4}$$

This is called the separating hyperplane. If denoted by ω which means w, the distance marked at the point x_i of the separating hyperplane (w, b) is given by::

$$d_i = \frac{w.x_i + b}{\omega} \tag{5}$$

Combination of the above inequalities and equations for all $x_i \in S$, then

$$y_i d_i \geq \frac{1}{\omega} \tag{6}$$

Therefore, $1/\omega$ is the lower bound on the distance between the x_i point and the separator hyperplane (w, b).
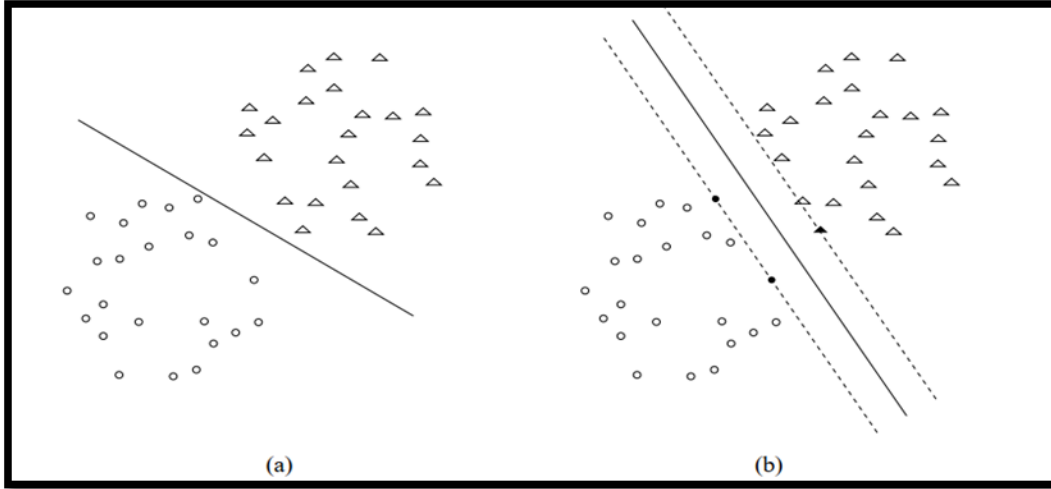

Figure 1. *Separating the Optimal Separator Hyperplane (a) and Hyperplane (b)*

Second, given the separating hyperplane (w,b) for the linear separable set S, the Canonical Representation of the separating hyperplane is obtained by changing the size of the pair (w,b) into pairs (w',b') such that the distance of the closest point equals 1/w'. Through this definition is obtained:

$$\min_{x_i \in S}\{y_i(w'.x_i + b)\} = 1 \tag{7}$$

Some of the commonly used Kernel functions are (Liu, 2018):

1. Linear kernel

$$K(x_i, x) = x_i^T.x \tag{8}$$

2. Polynomial kernel

$$K(x_i, x) = (\gamma.x_i^T.x + r)^d, \gamma > 0 \tag{9}$$

3. Radial Basic Function

$$K(x_i, x) = \exp(-\gamma|x_i - x|^2), \gamma > 0 \tag{10}$$

4. Sigmoid kernel

$$K(x_i, x) = \tanh(\gamma.x_i^T.x + r)^d, \gamma > 0 \tag{11}$$

In this case γ, r, and d are kernel parameters, and the C parameter is a penalty due to errors in the classification for each kernel.

2.1 SVM Parameters Optimization using PSO

In the beginning, PSO was created to solve problems for which the search space for solutions is continuous. PSO can also be used to search for solutions in a discrete search space. This can be done by a simple continuous multiple PSO method by rounding the continuous number to the nearest whole number. There are also more complicated ways to multiply continuous PSO to discrete PSO, for example by using the sigmoid function (Khanesar et al., 2007). Apart from continuous and discrete, PSO can also be used to solve combinatorial problems, for example to solve the problem of traveling salesmen (Wang et al., 2016). Tasgetiren et al. (2004) proposed a method for solving combinatorial problems using the smallest position value (SPV) method.

Parallel computing has been used to implement the PSO algorithm. Wang et al., (2016) and Mussi et al. (2009), implemented a multi-multiples PSO, namely Standard PSO (SPSO) using CUDA. Liera et al., (2011) implemented PSO to test the Rastrigin and Ackley functions in a 30-dimensional search space. Nashed et al., (2011) implemented PSO on GPU using OpenCL.

Some of the general terms used in PSO can be defined as follows (Tuegeh, et al., 2009):

1. Swarm: population of an algorithm.
2. Particle: member (individual) on a swarm.
   Each particle represents a potential solution to the problem being solved. The position of a particle is determined by the representation of the current solution.
3. Pbest (Personal best): Pbest position of a particle which indicates the position of the particle that is prepared to get the best solution.
4. Gbest (Global best): the best position of the particle on the swarm.
5. Velocity (vector): the vector that drives the optimization process which determines the direction in which a particle is needed to move to improve its original position.
6. Inertia weight: the inertia weight is symbolized by w, this parameter is used to control the impact of the velocity given by a particle.

III. Discussion

Preprocessing was carried out on the data, by handling the missing value. Missing value is information that is not available for an object (case). Missing value occurs because information for something about the object is not provided, it is difficult to find, or simply does not exist. Basically, missing value is not a problem for all data, especially if the amount is only a little, for example only 1% of all data. However, if the percentage of missing data is large enough, it is necessary to test whether the data that contains a lot of missing is still suitable for further processing or not. In the Wisconsin Breast Cancer Data (Original) Data Set which consists of 699 data, there are 19 data that have missing value in the 6th (sixth) attribute, namely the Single Ephitalial Cell Size attribute. In this study, the handling of missing values is done by removing data or objects that contain missing values.

After deducting data containing missing values, the data used in this study were 680 data. Search for the best SVM parameter values using PSO and perform classification with the support Vector Machine algorithm using Matlab. The problem solving procedure using SVM and PSO can be described in the following steps:

Step 1: Initialize SVM Parameter values. In the application of SVM, the SVM parameter value must be initialized first.

Step 2: Generate random initial particles. The SVM parameter generated in parallel as many as N particles.

Step 3: The SVM parameter values above are used to train the training data. Each particle is evaluated based on the fitness value. The fitness value in this study is the SVM classification accuracy value.

Step 4: Update the position and velocity of the particles. The PSO procedure requires the position and velocity of each particle. At this stage, the SVM model is again tested to find the best particles. The best particle evaluated based on the fitness value in step 4 is applied to retrain the training data and the fitness value is calculated.

Step 5: Update personal optimal fitness (*pfit*) and personal optimal position (*pbest*). So far all the particles in the initial (first) iteration have been evaluated, the first iteration produces the best particle (*pbest*).

Step 6: At this stage, evaluate whether all iterations have been completed. If it has not been continued in step 3, if it is finished then continue at step 8.

Step 7: Each iteration has the best particles. At this stage the best particle from all iterations will be evaluated to determine the best particle global optimal position (gbest).

Step 8: If the gbest value in step 7 is the expected value, then at that time the optimal SVM parameter value has been obtained and is continued in step 9. If the gbest value does not match the expected criteria, a new population will be generated again by returning to step 3. For more details, the classification process flow chart can be seen in Figure 4.
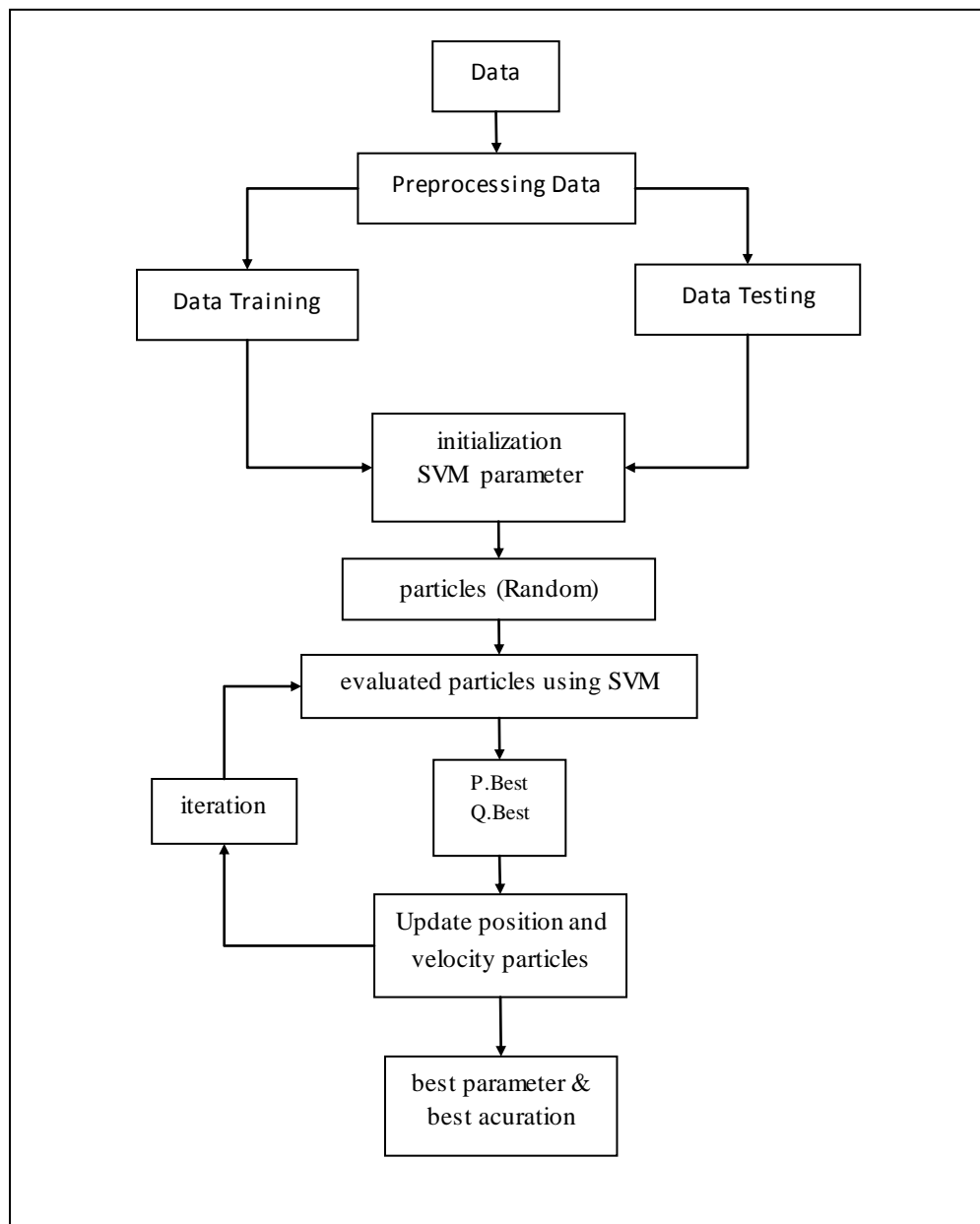
Figure 2. Research Framework

3.1 Classification Using SVM without Optimization

Each data was classified using a linear, polynomial, RBF, sigmoid kernel. The parameter values of each kernel are generated randomly in a certain range of values. Parameters c, γ, r are generated from 2^{-10} to 2^5 for linear, RBF, and sigmoid kernels, while the polynomial kernels are generated between 2^{-10} to 2^0. For parameter d (degree) it is generated randomly between 2 to 5.

Distribution of training data and test data on breast cancer data using the 10-fold cross validation method. The classification results of generating 30 randomly generated parameter pairs on breast cancer data in linear, polynomial, RBF and sigmoid kernels can be seen in Table 1.In Table 1 it can be seen that 30 random c parameter values are generated which are then tested into linear SVM, then record the results. classification accuracy. From the random selection of 30 c parameters, the lowest classification accuracy is 64.975% resulting from the selection of the parameter value c=0.001446, and the highest classification accuracy is 77.34% resulting from the selection of c parameter values from 20 to 30. Of the 30 parameter values c, γ, r and d generated randomly which are then tested on the polynomial SVM, then the classification accuracy results are recorded. From the random selection of 30 parameters c, γ, r, d, the lowest classification accuracy is 64.97% resulting from the selection of the 21st to 30th parameter values, and the highest classification accuracy is 77.08% resulting from the selection of the parameter value c=77.0868, γ=0.307749, r=0.10648 and d=0.967383. The results of the random selection of 30 parameters c and g for SVM RBF parameters obtained the lowest classification accuracy 64.71% resulting from the selection of parameter values c=9.369665 and γ=4.811239, and the highest classification accuracy is 78.26% resulting from the selection of values. parameter c=0.597232 and γ=0.067653. The results of the random selection of 30 parameters c, g and r for sigmoid SVM parameters obtained the lowest classification accuracy 59.76% resulting from the selection of parameter values c=0.802892, γ=1.050785 and r=2.507338, and the highest classification accuracy is 75.66% resulting from the selection of the parameter value c=0.026665, γ=7.553248 and r=2.901333.

Table 1. Classification 30 Parameters Using Linear, Polynomial, RBF and Sigmoid Kernel

	Linear			Polynomial						RBF				Sigmoid			
	Accuracy	C		Accuracy	C	γ	r	d		Accuracy	c	γ		Accuracy	c	γ	r
1	64.98	0.0014	1	64.98	0.3757	0.0041	0.0036	1	1	64.71	9.3697	4.8112	1	59.76	0.8029	1.0508	2.5073
2	65.10	0.0013	2	64.98	0.3544	0.0223	0.0037	3	2	65.10	0.0014	12.0806	2	61.33	0.4157	1.3604	5.0135
3	67.05	0.0024	3	65.10	0.0088	0.0025	0.0038	4	3	65.10	0.334	9.1566	3	64.07	18.0962	1.7602	0.0149
4	69.01	0.0027	4	65.10	0.0021	0.0032	0.0185	5	4	65.10	0.0071	0.4715	4	64.33	6.1857	15.9204	0.0305
5	74.87	0.0040	5	65.10	0.0021	0.0027	0.1527	5	5	65.10	0.0452	0.0012	5	64.59	1.1370	0.8177	0.0046
6	76.57	0.1152	6	65.10	0.0063	0.0265	0.0309	3	6	65.10	0.0050	0.0052	6	64.85	5.6324	24.1025	0.0805
7	76.57	0.2202	7	65.10	0.8273	0.0449	0.0116	0.5781	7	65.10	0.0488	0.0026	7	65.10	0.0086	0.2234	0.0016
8	76.70	0.0319	8	65.10	0.1344	0.0012	0.0016	0.2572	8	65.10	0.0071	0.4924	8	65.10	0.0114	0.0095	0.0122
9	76.70	0.1570	9	65.10	0.0085	0.0014	0.0587	2	9	65.10	0.1466	1.3559	9	65.10	0.0034	0.0029	0.0971
10	76.70	0.1575	10	65.10	0.5420	0.0139	0.0196	5	10	65.10	0.0034	1.4126	10	65.10	0.5394	0.0031	1.2457
11	76.70	0.0597	11	65.10	0.0057	0.0122	0.0013	2	11	65.10	0.0103	0.0014	11	65.10	0.1053	0.0019	0.0409
12	76.96	0.0150	12	65.10	0.0018	0.0391	0.6893	3	12	65.10	0.0533	0.002	12	65.10	0.1151	0.0655	2.0635
13	76.96	0.2876	13	65.10	0.0562	0.4103	0.0470	0.0834	13	65.10	0.0134	0.2437	13	65.10	0.0140	0.0072	0.0162
14	76.96	0.3999	14	65.10	0.6979	0.0095	0.0035	4	14	65.10	0.02	0.8807	14	65.10	0.2700	0.0026	0.0853
15	77.09	0.0149	15	65.10	0.0015	0.0885	0.0308	0.7125	15	65.10	0.0154	4.9236	15	65.10	8.2770	0.2302	9.9629
16	77.09	0.0445	16	65.10	0.0070	0.0017	0.9609	1	16	65.10	26.7219	23.0987	16	65.10	0.0034	0.1510	0.0604
17	77.22	0.0115	17	65.10	0.0037	0.0172	0.7706	3	17	65.10	0.0349	0.0287	17	65.10	17.1180	0.0584	4.3702
18	77.22	0.6421	18	65.10	0.0210	0.0083	0.1080	0.8480	18	65.1	0	0.5603	18	65.24	0.9527	0.6682	0.0970
19	77.22	2	19	65.10	0.0149	0.0061	0.0160	1	19	65.76	5.15383	1.7116	19	65.37	1.5575	29.9725	2.5036
20	77.34	0.0109	20	65.75	0.2317	0.0089	0.2194	2	20	67.05	0.42368	0.0029	20	65.49	0.7254	0.2790	0.4327
21	77.35	4	21	70.83	0.1795	0.0511	0.2540	4	21	68.23	0.44832	0.7464	21	66.54	0.1428	2.1718	0.0494
22	77.35	15	22	70.96	0.0074	0.2589	0.5366	3	22	73.44	0.08384	0.1307	22	66.66	0.0153	0.2424	0.0570
23	77.35	2	23	71.75	0.0371	0.5289	0.2661	5	23	76.31	1.93688	0.2449	23	67.05	0.5751	0.0041	0.0076
24	77.35	22	24	74.87	0.0562	0.8399	0.0354	2	24	76.56	0.11751	0.08127	24	67.06	0.8203	0.2814	0.0363
25	77.35	1	25	75.65	0.0224	0.1026	0.9484	3	25	77.35	26.4385	0.02522	25	67.85	0.7525	0.2178	0.0092
26	77.35	28	26	76.31	0.0383	0.3006	0.1879	5	26	77.48	7.12476	0.00627	26	68.36	0.0739	2.9891	0.0012
27	77.35	10	27	76.43	0.0821	0.2327	0.1758	3	27	77.48	0.41893	0.02709	27	70.19	2.9369	0.0438	1.1902
28	77.35	13	28	76.43	0.3259	0.1355	0.1944	4	28	77.61	1.02643	0.01469	28	73.44	0.0372	2.7370	1.4750
29	77.35	4	29	77.09	0.3030	0.0571	0.3656	2	29	77.87	0.79635	0.07929	29	73.70	0.9521	0.1033	0.0592
30	77.35	1	30	77.09	0.3077	0.1065	0.9674	2	30	78.26	0.59723	0.06765	30	75.66	0.0267	7.5532	2.9013

From the classification results of the generation of 30 randomly generated parameter pairs on breast cancer data in each kernel, the lowest and highest accuracy is shown in Table 2. In Table 2 it can be seen that the lowest accuracy is 59.75% obtained from the sigmoid kernel, while the highest is equal to 78.26% of the RBF kernel.

Table 2. Classification 30 Parameters

Kernel	Lowest accuracy (%)	Highest accuracy (%)
Linear	64.97	77.34
Polinomial	64.97	77.08
RBF	64.71	78.26
Sigmoid	59.75	75.66

3.2. Selection of the Best SVM Parameter Values Using PSO
a. Step-1: Initialization Parameter
This test aims to find the best SVM parameter value so as to maximize the accuracy of the classification of breast cancer data, which amounts to 680 data using the SVM classification. Prior to that, there were several parameters of the PSO and SVM parameters that were predefined, namely:

Number of Data	: 680
k-fold validation	: 10-fold.
Number of Training Data	: 612
Number of Testing Data	: 68
Number of particles	: 30 particles
inertia weight	: 1
cognitive weight	: 2
social weight	: 2
velocity max particle	: 10% from the max parameter

b. Step-2:
Generate the initial 50 particles with the initial parameter values. The initial value of the C parameter for the ith particle is $C_{i+1} = C_i + 0.2$, the initial value of the γ parameter for the ith particle is $\gamma_{i+1} = \gamma_i + 0.2$, the initial value of the r parameter for the ith particle is $r_{i+1} = r_i + 0.2$, the initial value of the d parameter in the ith particle is $d_{i+1} = d_i + 0.2$. The parameter values of c, γ, r, d will be used to train the training data. The variables c', γ', r', d' are the velocity vector values, which are parameters that affect the direction and velocity of the particle's movement, while the variables c", γ'', r", d" are the best positions for each particle. . If a particle gets the best accuracy (SVM accuracy according to its kernel after the current displacement is better than the previous displacement, then the values of c", γ'', r", d" are updated with the current position).

3.3 SVM Classification Using Optimization Method

Next, we will try to calculate the SVM classification combined with the PSO optimization method. The Matlab program used is the same when performing classification without optimization. The accuracy value is used as the objective function of the PSO. The initial parameter values that will be used as particles in the PSO are 30 parameters which produce the lowest accuracy for each data and each kernel. This study takes a low accuracy value as a particle because it also proves that PSO is able to optimize the level of accuracy even though it starts from a low value. The PSO will stop if the accuracy increases below 0.1% in 10 consecutive iterations. So, as long as the stop conditions are not met, PSO will continue to update these parameters. The results of the experiment on combining SVM and PSO using each kernel can be seen in Table 3.

In Table 3, the linear kernel section, the c values presented are the c values obtained using the PSO method. The value of c which is the initiation of the PSO process is the value of c which produces the lowest accuracy in Table 1. This is intended to compare the increase in SVM accuracy without PSO optimization with SVM accuracy with PSO optimization. In iteration 1 to iteration 5, there is an increase in accuracy from 65.10% to 77.74%. Then in the 5th iteration there is no longer an increase in accuracy. Iteration stops if there is no increase in accuracy in 10 iterations. In this test, the process stops at the 14th iteration with an accuracy of 77.74%.

Table 3. SVM-PSO Classification Using Linear, Polynomial, RBF and Sigmoid Kernel

Linear				Polynomial																
Iter	Accu	Difference	C	Iter	Accu	Difference	c	γ	r	d	Iter	Accu	Difference	C	γ	Iter	Accu	Difference	c	γ	r
1	65.1	65.1	0.0018	1	65.10	65.10	0.0389	0.0019	0.0905	2.0873	1	65.10	65.10	0.0070	1.0439	1	65.10	65.10	0.0535	0.0525	12.448
2	71.87	6.77	0.0032	2	69.40	4.29	0.3029	0.0102	0.4279	2.0268	2	65.23	0.13	0.1318	0.4003	2	73.96	8.86	0.0897	0.2721	0.0209
3	77.22	5.34	0.0071	3	78.00	8.60	1	0.0042	1	4.3369	3	77.34	12.11	0.2343	0.0135	3	76.43	2.47	0.1299	0.0382	0.0052
4	77.74	0.52	0.0122	4	78.00	0	1	0.0042	1	4.3369	4	77.74	0.39	1.3590	0.0109	4	77.09	0.65	0.3417	0.0440	0.0013
5	77.74	0	0.0122	5	78.12	0.13	1	0.0056	1	5	5	78.00	0.26	8.2747	0.0828	5	77.22	0.13	0.4518	0.0525	0.00098
6	77.74	0	0.0122	6	78.12	0	1	0.0071	1	4.9728	6	78.13	0.13	32	0.0476	6	77.35	0.13	0.6326	0.0202	0.00098
7	77.74	0	0.0122	7	78.13	0.002	1	0.0048	1	5	7	78.13	0	32	0.0476	7	77.48	0.13	0.8653	0.0119	0.00098
8	77.74	0	0.0122	8	78.13	0	1	0.0048	1	5	8	78.13	0	32	0.0476	8	77.61	0.13	2.3416	0.0041	0.00098
9	77.74	0	0.0122	9	78.13	0	1	0.0048	1	5	9	78.13	0	32	0.0476	9	77.61	0	2.3416	0.0041	0.00098
10	77.74	0	0.0122	10	78.13	0	1	0.0039	1	5	10	78.39	0.26	32	0.0501	10	77.61	0	2.3416	0.0041	0.00098
11	77.74	0	0.0122	11	78.26	0.13	0.1738	0.1206	1	4.7579	11	78.39	0	32	0.0501	11	77.61	0	2.3416	0.0041	0.00098
12	77.74	0	0.0123	12	78.39	0.13	0.2470	0.0711	1	4.7539	12	78.39	0	32	0.0501	12	77.61	0.001	2.1539	0.0057	0.00098
13	77.74	0	0.0123	13	78.39	0	0.2470	0.0711	1	4.7539	13	78.39	0	32	0.0501	13	77.61	0	2.1539	0.0057	0.00098
14	77.74	0	0.0123	14	78.39	0	0.2470	0.0711	1	4.7539	14	78.39	0	32	0.0501	14	77.61	0	2.1539	0.0057	0.00098
				15	78.65	0.26	0.2503	0.0556	1	5	15	78.39	0	32	0.0501	15	77.61	0	2.1539	0.0057	0.00098
				16	78.65	0	0.2503	0.0556	1	5	16	78.39	0	32	0.0501	16	77.61	0	2.1539	0.0057	0.00098
				17	78.65	0	0.2503	0.0556	1	5	17	78.40	0.001	22.574	0.0603	17	77.61	0	2.1539	0.0057	0.00098
				18	78.65	0	0.2503	0.0556	1	5	18	78.52	0.13	20.646	0.0576	18	77.61	0	4.0084	0.0031	0.00098
				19	78.65	0	0.2503	0.0556	1	5	19	78.53	0	20.646	0.0576						
				20	78.65	0	0.2503	0.0556	1	5	20	78.53	0	13.997	0.0700						
				21	78.65	0	0.2503	0.0556	1	5	21	78.53	0	13.997	0.0700						
				22	78.91	0.26	0.2521	0.0532	1	5	22	78.53	0	13.997	0.0700						
				23	78.91	0	0.2521	0.0532	1	5	23	78.53	0	13.997	0.0700						
				24	78.91	0	0.2521	0.0532	1	5	24	78.53	0	13.997	0.0700						
				25	78.91	0	0.2521	0.0532	1	5	25	78.53	0	13.997	0.0700						
				26	78.91	0	0.2521	0.0532	1	5	26	78.53	0	13.997	0.0700						
				27	78.91	0	0.2521	0.0532	1	5	27	78.53	0	13.997	0.0700						
				28	78.91	0	0.2521	0.0532	1	5	28	78.53	0	14.580	0.0689						
				29	78.91	0	0.2521	0.0532	1	5											
				30	78.91	0	0.2521	0.0532	1	5											
				31	78.91	0	0.2521	0.0532	1	5											
				32	78.91	0	0.2521	0.0532	1	5											

In the polynomial kernel, the values of the SVM parameters c, γ, r and d presented are the values obtained using the PSO method. The parameter values c, γ, r, and d which are the initiations of the PSO process in the polynomial kernel are the values that produce the lowest accuracy in Table 1. This is intended to be compared to the increase in SVM accuracy without PSO optimization with SVM accuracy with PSO optimization. There have been several improvements in iterations and at the 21st iteration there has been no improvement in accuracy. However, in the 22nd iteration there was still an increase of 0.26% and then there was no longer an increase in accuracy. The process stops at the 32nd iteration after there is no increase in accuracy from 23 to 32 (10 iterations), and the result is an accuracy of 78.91%.

In the RBF kernel, the values of the SVM c and γ parameters presented are the values obtained using the PSO method. The parameter values c and γ which initiate the PSO process in the polynomial kernel are the values that produce the lowest accuracy in Table 4.4. This is intended to compare the increase in SVM accuracy without PSO optimization with SVM accuracy with PSO optimization in the RBF kernel. The process stops at the 28th iteration after there is no increase in accuracy from the 19th to 28th iterations (10 iterations), and the highest accuracy results are 78.52%.

In the sigmoid kernel, the values of the SVM c, γ and r parameters presented are the values obtained using the PSO method. The parameter values of c, g and r which are the initiations of the PSO process in the polynomial kernel are the values that produce the lowest accuracy in Table 1. This is intended to be able to compare the increase in SVM accuracy without PSO optimization with SVM accuracy with PSO optimization on the Sigmoid kernel. The process stops at the 18th iteration. In the 12th iteration there is an increase in accuracy by 0.000849, but because it is smaller than 0.001% it is ignored and the iteration is continued. After there was no increase in accuracy from the 19th to 28th iterations which was more than 0.01%, the highest accuracy result was 78.52%. Comparison of experimental results between the highest accuracy SVM classification without PSO and the highest accuracy SVM classification with PSO optimization can be seen in Table 4.

Table 4. SVM-PSO Classification

Kernel	SVM Accuration	SVM-PSO Accuration	Improved Accuracy	Iteration	Time (s)
Linear	65.10 %	77.73 %	12.63 %	14	53.3562
Poli	65.10 %	78.91 %	13.80 %	31	275.982
RBF	65.10 %	78.52 %	13.42 %	28	296.391
Signoid	65.10 %	77.60 %	12.50%	18	271.362

In Table 4, it can be seen that PSO can improve accuracy even though it starts from a small accuracy. This proves that PSO is effective for finding parameter values that can produce high accuracy. The results of increasing the accuracy of each data in each kernel are also shown in Table 4. When compared to the classification results without the optimization method using the optimization method, the SVM classification results using PSO optimization are better than without PSO. The comparison of results without the optimization method using the PSO optimization method is presented in Table 5.

Table 5. The Comparison of Results without Optimization Method Using Optimization Method

Kernel	SVM	SVM-PSO
Linear	77,34	77,73
Polinom	77,08	78,91
RBF	78,26	78,52
Sigmoid	75,66	77,60

From Table 5, the accuracy of SVM using PSO is better than the results of SVM classification with random parameter search. This shows that the search for parameters using an unsystematic way allows us to get stuck on non-optimal results, whereas when using a more systematic way, namely using an optimization method such as PSO in finding parameter values, SVM is able to produce more optimal accuracy. Each iteration process has a different execution time, depending on the number of instances and parameter values of each kernel. The execution time of each classification process with optimization and execution time without optimization is presented in Table 6. Apart from being influenced by the number of instances, the execution time is also influenced by the magnitude of the parameter value. The greater the parameter value, the longer the execution time.

Table 6. SVM-PSO Classification Timing Processing

Kernel	SVM-PSO		SVM
	Time (s)	Number of Iterations	Time (s)
Linear	53,3562	14	12,783
Polinom	275,9815	32	5,832
RBF	296,3906	28	11,345
Sigmoid	271,3617	18	15,957

The optimal classification accuracy when using the polynomial kernel with an accuracy of 78.91% and the parameter values of c, γ, r, and d are 0.252101, 0.053248, 1 and 5, respectively which is higher than the linear and sigmoid kernels. However, it cannot be ascertained that the polynomial and RBF kernel are the best kernels because the kernel selection depends on the data to be used.

IV. Conclusion

Each kernel in the SVM procedure has parameters that can confirm the results of the data classification. It is not easy to ascertain the parameter value of each kernel, so in this research SVM is combined with PSO optimization procedures to find the best parameter value. The combination of SVM and PSO is quite efficient in finding the maximum parameter value. This combination produces a method that is more systematic than the trial and error method in finding the SVM parameter value. However, this merge requires a longer execution time to reach the maximum value. The use of SVM in classification with optimization using PSO has been able to improve accuracy better than the classification using SVM without optimization, where the determination of the parameters is done randomly. The classification accuracy of the breast cancer data can be increased to 78.91%.

In this research, we did not try to change the inertia weight value and the correction factor. Both of these aspects will affect the velocity value of the PSO particles. In the next research, it is hoped that the use of a portion of the inertia weight value and the acceleration coefficient is used to find the best value. Not only that, we can also use dynamic population in PSO. If one day the particles are trapped at something local optimum, these particles can be removed or exchanged with new particles.

References

Arysha, V.O., Santosa, H., and Sanusi, S.R. (2020). The Effect of Extrinsic Motivation on Housewife Actionsin IVA Testin the Working Area of Community Health Center at Bandar Kalipah. Britain International of Exact Sciences (BIoEx) Journal Vol. 2 (1):352-356.

Demidova, L, Klyueva, I, & Pylkin, A 2016, 'The study of characteristics of the hybrid particle swarm algorithm in solution of the global optimization problem', *2016 5th Mediterranean …,,* ieeexplore.ieee.org

Han JH, et al. 2011, *Effect of glucose on xylose utilization in Saccharomyces cerevisiae harboring the xylose reductase gene. Arch Microbiol, Vol.*193(5), pp: 335-40

Khasenar et., al 2007, Sliding mode control of Rotary Inverted Pendulm, Conference: Control & Automation, 2007, July 2007.

Liera, IC, Liera, MAC, & Castro, M 2011, 'Parallel particle swarm optimization using GPGPU', *… de Las Villas 14 al 18 …,,* pdfs.semanticscholar.org

Liu, N. *et al.* (2018) 'Improved Cost-Sensitive Support Vector Machine Classifier for Breast Cancer Diagnosis', 2018.

Mussi, L, Cagnoni, S, & Daolio, F 2009, 'GPU-based road sign detection using particle swarm optimization', *2009 Ninth International …,,* ieeexplore.ieee.org

Nashed, A, Saikia, P, Herrmann, WA, Gabel, VP, & ... 2011, 'The outcome of early surgical repair with vitrectomy and silicone oil in open-globe injuries with retinal detachment', *American journal of …,,* Elsevier

Pontil, M, & Verri, A 1997, 'Proprieties of support vector machines', *Relatório técnico, Massachusetts Institute of …,*

Rachman, F, & Purnami, SW 2012, 'Perbandingan klasifikasi tingkat keganasan breast cancer dengan menggunakan regresi logistik ordinal dan support vector machine (SVM)', *Jurnal sains dan seni ITS,,* ejurnal.its.ac.id

Salama, G. I., Abdelhalim, M. B. and Zeid, M. A. (2012) 'Experimental comparison of classifiers for breast cancer diagnosis Experimental Comparison of Classifiers for Breast Cancer Diagnosis', (January 2016). doi: 10.1109/ICCES.2012.6408508.

Tasgetiren, MF, Liang, YC, Sevkli, M, & ... 2004, 'Particle swarm optimization algorithm for makespan and maximum lateness minimization in permutation flowshop sequencing problem', *Proceedings of the …,,* researchgate.net

Tuegeh, M., Soeprijanto and Purnomo, M. H. (2009) 'Modified improved particle swarm optimization for optimal', *Seminar Nasional Aplikasi Teknologi Informassi 2009 (SNATI 2009)*, 2009(Snati), pp. 85–90.

Wang, C, Lin, M, Zhong, Y, & ... 2016, 'Swarm simulated annealing algorithm with knowledge-based sampling for travelling salesman problem', *International Journal of …,,* inderscienceonline.com.