Analyzing cerebral infarction using support vector machine with artificial bee colony and particle swarm optimization feature selection

Z Rustam\(^1\), D A Utami\(^2\), J Pandelaki\(^3\), and R E Yunus\(^4\)

\(^1,2,3,4\)Department of Mathematics, Universitas Indonesia Depok, 16424, Indonesia

Abstract. Early diagnosis of cerebral infarction is essential since many patients cannot be cured where the diagnosis is made at an advanced stage. In case an infarct occurs, the tissue in the brain die and stop the circulation of blood, which carries oxygen and nutrients to the body. Therefore, this study uses a machine learning Support Vector Machine (SVM) for early detection of the disorder. To produce the best classification accuracy and fast computing time, feature selection is performed on cerebral infarction data, including Artificial Bee Colony (ABC) and Particle Swarm Optimization (PSO). After classification, infarction data with the best features are classified using SVM. The classification results of ABC-SVM and PSO-SVM methods are compared with the accuracy of 90.36% for ABC-SVM and 86.74% for PSO-SVM. Therefore, the best approach used in classification is the SVM method with ABC feature selection.

1. Introduction

Stroke is a condition which occurs due to sudden disruption of blood flow to the brain [1]. It is caused by a blockage called an infarction or rupture of blood vessels. Generally, blood clots heart and other large blood vessels are the leading causes of blood vessel infarction [1]. If a stroke occurs, the tissue in the brain die and stop the circulation of blood, which carries oxygen and nutrients to the body [1]. In human, there are 100 billion nerve cells and trillions of nerve connections in the brain. Basically, 70% of the oxygen and nutrients needed by the human body are used to ensure the brain functions properly [2]. Unlike the muscles, the brain is unable to store nutrients as reserves, and therefore, there is a need for blood flow in order for it to work appropriately.

In general, stroke is divided into two type, hemorrhagic and ischemic stroke [3]. Of all stroke cases, two-thirds are ischemic, and the rest are hemorrhagic. The hemorrhagic is caused by the occurrence of high blood pressure or other diseases which weaken the blood vessels, making them burst [3]. The blockage of blood vessels causes ischemic stroke due to blood clots in the heart, and hardening of the arteries, a condition referred to as atherosclerosis [3]. This condition is referred to as cerebral infarction, and it is more common in ischemic stroke. Cerebral infarction is a condition of tissue damage in the brain resulting from the insufficient oxygen supply. Generally, the insufficiency is brought about due to the disruption of blood flow to the brain [4]. Damage and death of the tissues in the brain result in decreased neurological function.

The presence of cerebral infarction in the brain helps in the diagnosis of the ischemic stroke, mainly detected by conducting a CT scan. However, the results obtained is not enough to detect and diagnose early cerebral infarction. For this reason, machine learning is often used in the early classification of this
condition, especially in cerebral infarction using labels and features available from the results of CT scans.

The machine learning used for classification in this study is Support Vector Machine (SVM). SVM works based on structural risk minimization (SRM) principles or by abating basic risk to obtain the best plane (hyperplane) for separating two classes in the input space [5]. Before using SVM, cerebral infarction data are selected through Artificial Bee Colony (ABC) and Particle Swarm Optimization (PSO) feature selection. The selection approach is used to eliminate redundant features and increase the accuracy value. Feature selection method shows the level of relevance of a feature if the weight is positive. However, the weight value close to zero and negative means a feature is not relevant for use. After the data features are selected, the classification is carried out using SVM. The results of the classification of cerebral infarction using ABC-SVM and PSO-SVM should be compared.

There are several studies discussing SVM and feature selection. Rustam and Yaurita [6] predict the insolvency using SVM and Fuzzy Kernel C-Means (FKCM) method. The result show that SVM and FKCM can be a useful tool for parties who are interest in evaluating insolveny of an insurance firm. Panca and Rustam [7] improved to solve multiclass classification using a method based on support vector machine recursive feature elimination (SVM-RFE) and Twin SVM. This method is used as the classifier to reduce computational complexity and to find two non-parallel optimum hyperplanes.

Another study proposed a new method of hybrids based on the correlation feature selection method and the artificial bee colony algorithm, Co-ABC to select a small number of relevant genes for accurate classification of gene expression profiles [8]. This study proves that Co-AbC is an efficient approach to the discovery of biomarker genes using cancer genetic factor expression profiles.

A research propose a method to integrate Correlation-based Feature Selection (CFS) with Binary Particle Swarm Optimization (iBPSO) [9]. This model selects a set of low-dimensional prognostic genes to classify biological samples from binary and multi-class cancers using the Naive-Bayes classifier. The proposed IBPSO also controls the convergence problem. The model shows better results in terms of classification accuracy and the number of genes selected in many cases.

2. Material and Methods

2.1. Cerebral Infarction Dataset

Data were obtained from the Department of Radiology, Dr. Cipto Mangunkusumo (RSCM) Hospital. The statistics used for cerebral infarction was specifically from ischemic stroke patients between July and November 2018. This involved 206 cerebral infarction cases with seven basic features, classified as 30% testing and 70% of training. Table 1 shows an example of cerebral infarction data, while Table 2 explains the seven features.

Cerebral infarction data are divided into two classes. This includes class 0, which involves 103 normal patients with no cerebral infarction in the brain. The second category was class 1, with 103 positive patients with cerebral infarction in the brain.
Table 1. Example of Cerebral Infarction Dataset

| Area (cm²) | Min | Max | Average | SD  | Sum | Length (cm) | Target |
|------------|-----|-----|---------|-----|-----|-------------|--------|
| 0          | 33  | 38  | 35.38   | 1.92| 283 | 0.5         | 0      |
| 0.1        | 22  | 46  | 31.31   | 5.47| 1221| 1.2         | 0      |
| 0.1        | 18  | 51  | 31.41   | 9.33| 4146| 1.6         | 0      |
| 0.1        | 12  | 50  | 31.36   | 9.78| 7934| 1.6         | 0      |
| 0.1        | 15  | 52  | 32.4    | 8.54| 5411| 1.4         | 0      |
| 0.1        | 21  | 51  | 34.67   | 6.37| 3432| 1.6         | 0      |
| 0.1        | 25  | 59  | 44.73   | 6.16| 3847| 1.5         | 0      |
| 0.3        | -15 | 38  | 11.16   | 9.36| 3829| 2.4         | 1      |
| 0          | 5   | 28  | 15.36   | 6.21| 553 | 1.1         | 1      |
| 0          | -13 | 14  | 1.86    | 7.71| 52  | 0.9         | 1      |

Table 2. Feature of Cerebral Infarction

| No. | Feature | Definition of Feature |
|-----|---------|-----------------------|
| 1.  | Area    | Area of infarct       |
| 2.  | Min     | The minimum value of infarct |
| 3.  | Max     | The maximum value of infarct |
| 4.  | Average | The average value of infarct |
| 5.  | SD      | Standard error value of infarct |
| 6.  | Sum     | Sum of acute point of infarct |
| 7.  | Length  | Length of infarct     |

2.2. Artificial Bee Colony (ABC) Feature Selection

Karaboga made the ABC algorithm by simulating the way bees collected nectar in 2005. Optimal results are obtained through the collection of nectar [10].

In the ABC algorithm, $x_i$ solution is initialized in space with D-dimension with iteration $i = 1, 2, ..., N$. The system involves one bee that attracts one food source. The position of the food source is the distance from the bee employed [10]. The location is determined based on the probability of the bee selecting its food source.

$$F_i = \frac{f(t_i)}{\sum_{n=1}^{N} f(t_n)} \quad (1)$$
Where \( F_i \) states the probability of selection and \( fit_i \) the fitness of \( i \)-food source. The observer bee conducts an environmental search using the formula:

\[
x_{i}^{k+1} = x_{i}^{k} + r \times (x_{j}^{k} - x_{i}^{k}) \tag{2}
\]

Where \( i \neq j \) with \( j = 1, 2, ..., N \) and \( x_{i}^{k} \) represent the location of the food source, and \( r \) denotes the number randomly distributed in the interval (-1,1). Then, the value of fitness before and after the bee searching for the food and environment, the chosen one has the best solution [11]. The formula for producing new food sources, which is

\[
x_i = x_{\text{min}} + \text{rand}(0,1)(x_{\text{max}} - x_{\text{min}}) \tag{3}
\]

2.3. Particle Swarm Optimization (PSO) Feature Selection

The PSO algorithm comes from the study of bird predation behavior, first proposed by Kennedy and Eberhart in 1995[2]. This algorithm produced a sound output [12].

Suppose that in D-dimensional space, the population of particle \( n \) is \( X = [X_1, X_2, ..., X_n] \). The position of the \( i \)-particle is \( X_i = [x_{i1}, x_{i2}, ..., x_{id}] \), and the \( i \)-particle velocity is \( V_i = [v_{i1}, v_{i2}, ..., v_{id}] \) determines the direction and distance of particle motion. These specks move in a D-dimensional solution space. In this case, the best particle position should be tracked, \( B_i \) best personal position, and \( B_g \) best global position. This particle updates its position and speed through the best position [12]. The best personal position is the optimum location for particles individually [12], namely \( B_i = [b_{i1}, b_{i2}, ..., b_{id}] \) while the best global position is the most appropriate spot of all particles in the population [12], which is \( B_g = [b_{g1}, b_{g2}, ..., b_{gd}] \). The following is the formula [13]:

\[
v_{id}^{k+1} = \omega v_{id}^{k} + c_1 r_1 (b_{id}^{k} - x_{id}^{k}) + c_2 r_2 (b_{gd}^{k} - x_{gd}^{k});
\]

\[
d = 1, 2, ..., D \tag{4}
\]

\[
x_{id}^{k+1} = x_{id}^{k} + v_{id}^{k+1}; d = 1, 2, ..., D; \tag{5}
\]

\[
i = 1, 2, ..., n
\]

Where \( k \) denotes the number of iterations, \( \omega \) the weight of inertia, \( c_1 \) and \( c_2 \) show cognitive and social learning factors, \( r_1 \) and \( r_2 \) state that the numbers are uniformly distributed in intervals (0.1). Speed and position of each particle have a boundary interval expressed as \([V_{\text{min}}, V_{\text{max}}]\) and\([X_{\text{min}}, X_{\text{max}}]\).

2.4. Support Vector Machine

SVM works based on the principle of structural risk minimization (SRM) or by minimizing structural risk to obtain the best plane (hyperplane), which separate two classes in the input space [5]. The best hyperplane as a separator between the two classes might be found by calculating the margin of the hyperplane and determining the maximum point of the margin. The margin is the distance between the hyperplane and the closest data from each class. The data which is closest to the hyperplane is called the support vector [14].

For example given a dataset notated as \( x_i \in R^n \) which has a label \( y_i \in \{ -1, +1 \} \) for \( i = 1, 2, ..., n \) where \( n \) is the amount of data, the formula for hyperplane between class -1 and +1 if it is assumed a hyperplane completely separates the two classes with dimension \( n \) is as follows

\[
H_i = w^T x_i + b = 0
\]
$H_1$ represents the hyperplane, $w$ is the normal plane and $b$ the bias in the optimal hyperplane. The optimal hyperplane function is obtained by finding the weight parameter $w$ and the bias factor $b$ in the decision function below [5].

$$f(x) = w^T x + b$$

The $x_i$ data included in the negative class should be formulated as data that meets the equation below [15]

$$w^T x_i + b \leq -1, \text{ untuk } y_i = -1, \ i = 1, 2, ..., n$$

The $x_i$ data included in the positive class should be formulated as data that meets the equation below

$$w^T x_i + b \geq +1, \text{ untuk } y_i = +1, \ i = 1, 2, ..., n$$

Therefore, for each data, the following inequality forms apply.

$$y_i(w^T x_i + b) \geq 1, \ i = 1, 2, ..., n$$

The margin of the hyperplane is the distance between the $d_+$ dan $d_-$ fields. The problem of maximizing the margin $|d_+ - d_-|$ is equivalent to minimizing the value of $\|w\|$, for it to be written as a primal optimization problem as below

$$\min \frac{1}{2} \|w\|^2$$

Subject to

$$y_i(w^T x_i + b) \geq 1, \ i = 1, 2, ..., n$$

In solving the primal optimization problem, a solution might be used by changing it into a dual form using the Lagrange multiplier [15]

$$\min L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^{n} \alpha_i [y_i(w^T x_i + b) - 1]$$

Subject to

$$\alpha_i \geq 0$$

$$1 - y_i(w^T x_i + b) \geq 0$$

$$\alpha_i [1 - y_i(w^T x_i + b)] = 0$$

$$\forall i = 1, 2, ..., n$$

The solution to this problem takes the following form
The solution to this problem is the points with values of $\alpha_i \neq 0$, which are referred to as support vectors and are located in one of the hyperplanes which meet the values $f(x_i) = 1$ or $f(x_i) = -1$. For other points has a value of $\alpha_i = 0$ and lies in the positive class or negative class [14].

### 3. Proposed Method

This study used the Support Vector Machine (SVM) method in the classification of cerebral infarction in ischemic stroke patients. Before being classified using SVM, the feature of cerebral infarction dataset was filtered. This was meant to find out which features are very influential and features not relevant. The feature selection used is the Artificial Bee Colony (ABC) and Particle Swarm Optimization (PSO). Afterward, the ABC-SVM and PSO-SVM methods should be compared based on accuracy, precision, recall, and f1 score value. Figure 1 shows an illustration of the method.

**Figure 1.** an illustration of the Method

### 4. Results

#### 4.1. Result of Artificial Bee Colony (ABC) Feature Selection

Feature selection was performed on cerebral infarction data using ABC. The main goal was to find the features providing the most information to the one with the least info or no effect. The trick is to compare through the results of the weights of each feature from highest to lowest. The weight value results are in the interval 0 to 1. The highest value indicates the most optimal features, while the lowest one is not too influential in classifying data. Table 3 displays the weight values obtained from feature selection on cerebral infarction data using ABC Algorithm.
Table 3. Feature Score of Artificial Bee Colony

| Feature Score | Area 0.0155655269811 |
|---------------|---------------------|
| ABC           |                     |
| Min           | 0.300813716013      |
| Max           | 0.16167790009       |
| Average       | 0.42045372541       |
| SD            | 0.0190895910639     |
| Sum           | 0.0477101779394     |
| Length        | 0.0240949271898     |

In Table 3, the result of the largest weight value is 0.42045372541 from the Average feature, and the smallest is 0.0155655269811 from the Area feature. It is sorted according to the highest weight value to lowest, and the order is Average, Min, Max, Sum, Length, SD, and Area.

4.2. Result of Particle Swarm Optimization (PSO) Feature Selection

Table 4 displays the weight values obtained from feature selection on cerebral infarction data using PSO Algorithm.

Table 4. Feature Score of Particle Swarm Optimization

| Feature Score | Area 0.00742115027829 |
|---------------|---------------------|
| PSO           |                     |
| Min           | 0.325668559435      |
| Max           | 0.161137386008      |
| Average       | 0.439691758339      |
| SD            | 0.0306186462802     |
| Sum           | 0.0169891860777     |
| Length        | -0.0002857892468    |

In Table 4, the result of the largest weight value is 0.439691758339 from the Average, and the smallest is -0.0002857892468 from the Length feature. In case it is sorted from the highest weight value to the lowest, the order is Average, Min, Max, SD, Sum, Area, and Length.
4.3. Result of SVM with Artificial Bee Colony (ABC) and Particle Swarm Optimization (PSO) Feature Selection Method

The first four features should be taken with the largest value from ABC and the PSO method. For ABC selection, the features used are Average, Min, Max, and Sum. The PSO feature selection uses Average, Min, Max, and SD. Table 5 and Figure 2 are the comparisons of the results of the ABC-SVM and PSO-SVM classification report.

Table 5. Classification Report of the Method

| Method    | Accuracy | Precision | Recall  | f1 score |
|-----------|----------|-----------|---------|----------|
| ABC-SVM   | 90.36%   | 86.84%    | 91.60%  | 89.15%   |
| PSO-SVM   | 86.74%   | 90%       | 83.72%  | 86.75%   |

![Classification Report](image)

Figure 2. Classification Report of the Method

From Table 5 and Figure 2, the accuracy value of the ABC-SVM method has a better value than the PSO-SVM, precisely 90.36%, and 86.74%. Based on the accuracy, ABC-SVM is a better method compared to PSO-SVM. However, based on the precision value, the PSO-SVM is better than ABC-SVM since it is 90% while ABC-SVM is 86.84%. Additionally, based on the recall value, the ABC-SVM method is better compared to PSO-SVM with values of 91.60% and 83.72%. Looking at the f1-score value, the ABC-SVM method is better than the PSO method with a score of 89.15% compared to 86.75%.

5. Conclusion

This study classified cerebral infarction data using the Support Vector Machine (SVM) method and selection of Artificial Bee Colony (ABC) and Particle Swarm Optimization (PSO) features. After selection features using ABC and PSO methods in cerebral infarction data, the best four features were selected from each of these approaches. Afterward, the data with four features were classified using SVM and the accuracy, recall, precision, and f1 score of the ABC-SVM and PSO-SVM methods calculated. The technique with the most excellent accuracy, recall, precision, and f1 score is the best in the classification of cerebral infarction. Based on the results, the ABC-SVM method is the best compared to
the PSO-SVM since its accuracy, recall, and f1 scores are high, 90.36%, 91.60%, and 89.15 % respectively.

Acknowledgments
This research supported financially by the Ministry of Research and Higher Education Republic of Indonesia with PDUPUT 2019 research grant scheme, ID number 1621/UN2.R3.1/HKP.05.00/2019.

References
[1] Wang, G., Jing, J., Pan, Y., Meng, X., Zhao, X., Liu, L., Li, H., Wang, D., & Wang, Y 2019 Does All Single Infarction Have Lower Risk of Stroke Recurrence Than Multiple Infarctions in Minor Stroke?. BMC Neurology, 19, 7
[2] Omar, H 2011 Myocardial Infarction Stroke Association International Journal of Cardiology, 154(3), 340
[3] Mentari, I., Naufalina, R., Rahmadi, M., Khotib, J 2018 Development of Ischemic Stroke Model by Right Unilateral Common Carotid Artery Occlusion (RUCCAO) Method Fol Med Indones, 54(3), 200-206
[4] Zhao, Lei., Biesbroek, J., Shi, L., Liu, W., Kujif, H., Chu, W., Abrigo, J., Lee, R., Leung, T., Lau, A., Biessels, G., Mok, V., & Wong, A 2017 Strategic Infarct Location for Post-Stroke Cognitive Impairment: A Multivariate Lesion-Symptom Mapping Study. Journal of Cerebral Blood Flow & Metabolism, 38(8), 1299-1311
[5] Xue, H., Yang, Q., & Chen, S 2009 SVM: Support Vector Machines United States of America: Taylor & Francis Group
[6] Rustam,Z., Yaurita,F 2018 Insolvency Prediction in Insurance Companies Using Support Vector Machines and Fuzzy Kernel C-Means Journal of Physics: Conference Series Vol.1028(1)
[7] Panca,V., Rustam,Z 2016 Application of Machine Learning on Brain Cancer Multiclass Classification AIP Conference Proceedings, Vol.1862
[8] Hala, M 2018 Co-ABC: Correlation Artificial Bee Colony Algorithm for Biomarker Gene Discovery using gene Expression Profile Saudi Journal of Biological Sciences 25, 895-903
[9] Indu, J., Vinod, K., Renu, J 2018 Correlation Feature Selection Based Improved-Binary Particle Swarm Optimization for Gene Selection and Cancer Classification Applied Soft Computing Vol. 62, 203-215
[10] Beatriz, A., Katya, R., Roberto, V 2016 Classification of DNA Microarrays using Artificial Neural Networks and ABC Algorithm Applied Soft Computing 38, 548-560
[11] Ezgi, Z., Selma, A 2016 A Hybrid Approach of Differential Evolution and Artificial Bee Colony for Feature Selection Expert Systems With Applications 62, 91-103
[12] Laith, M., Ahamad, T 2017 Unsupervised Text Feature Selection Technique Based on Hybrid Particle Swarm Optimization Algorithm with Genetic Operators for the Text Clustering J Supercomput 73, 4773-4795
[13] Ali, A., Ali B 2018 Image Steganalysis using Improved Particle Swarm Optimization Based Feature Selection Appl Intell 48, 1609-1622
[14] Wang, L 2005 Support Vector Machines: Theory and Applications. New York: Springer, 177
[15] Farquad, M., Bose, I 2012 Preprocessing Unbalanced Data Using Support Vector Machine. Decision Support Systems, 53, 226-233