Acute sinusitis classification using support and fuzzy support vector machines

Z Rustam$^{1, *}$, N Angie$^1$, J Pandelaki$^2$, R E Yunus$^2$

$^1$Department of Mathematics, University of Indonesia, Depok 16424, Indonesia
$^2$Department of Radiology, Cipto Mangunkusumo Hospital, Jakarta 10430, Indonesia

$^*$Corresponding author: nurliaangied@sci.ui.ac.id, *rustam@ui.ac.id

Abstract. The medical sector is currently in need of a method to aid in the classification of diseases, which contemporarily progresses into varying types. Therefore, the role of technology is highly relevant in the process of overcoming this challenge. This report discusses acute sinusitis, which is one of the most common forms of sinusitis, possibly caused by viruses, bacteria, fungi, pollutants, allergies, and also autoimmune reactions. Furthermore, the Support Vector Machines (SVM) and Fuzzy Support Vector Machines (FSVM) are used as a classification method to diagnose a person of acute sinusitis, therefore, this research aims to compare how both work, using Radial Basis Function (RBF) and Polynomial Kernel. Data of CT scan from Cipto Mangunkusumo Hospital, Indonesia was used to evaluate acute sinusitis, in terms of Accuracy, Sensitivity, Precision, and F1-Score. Thus, the final results indicate a better performance for FSVM than SVM in all perspectives, especially using the RBF kernel.

1. Introduction

Acute sinusitis is a major manifestation of sinusitis, which causes inflammation and swelling of spaces inside the nose [1]. Therefore, a sufferer possesses the tendency of difficulty in breathing [1], and also presents with other symptoms, including facial pain or a headache and also swelling in areas around the eyes or face [1].

This is most often caused by the common cold, a viral infection [1], although there are other infrequent reasons, which include cystic fibrosis, neoplasia, and mechanical ventilation [2]. Moreover, individuals with sinusitis exhibit numerous symptoms or signs, encompassing fever, cough, hyposmia, nasal congestion, nasal drainage, fatigue, maxillary dental pain, postnasal drip, facial pain, and ear pressure [2].

The methods used in diagnosis comprises of nasal endoscopy, Computed Tomograph Scanning (CT scan), Magnetic Resonance Imaging (MRI), nasal and sinus cultures, and allergy test [3]. These provide detailed description of the condition [3], and antibiotics are often required on instances where it is caused by bacterial infection, in order to prevent the disease proliferation [3].

This research uses the Support Vector Machines and the Fuzzy variety, which is an extended technique, as methods of diagnosing a patient of acute sinusitis. Therefore, they are conducted with the
expectation of rendering assistance to medical staffs, in order to enhance the ease of diagnosis, as well as its efficiency.

Previous researches using other methods of classification include Kernel Based Fuzzy C Means, Kernel Spherical K-Means, $X^2$ Test and Binary Logistic Regression [4], Automatic Localization [5], Imaging Features [6, 7, 8], Cancer Classification [9], Brain Cancer Multiclass [10], and High-Dimensional Breast Cancer Database [11]. In addition, the Fuzzy Support Vector Machines was already being used in The Prediction of Bank Failures [12] and Class Imbalance Learning [13]. Meanwhile, the Support variety is currently applied in Detection Systems for Intrusions [14], Cancer Classification [15, 16], Detection of Traffic Incident [17], Sorting of Hyper Spectral Imagery [18], Schizophrenia Classification [19], Face Recognition [20, 21], Analysis of Gene Expression Data [22], and Insolvency Prediction [23].

2. Methods

2.1 Data
This research employed the use of a dataset from CT scan of patients suffering from acute sinusitis at the Department of Radiology, Cipto Mangunkusumo Hospital, Jakarta, Indonesia, comprising of four features, which include Gender, Age, Hounsfield Unit (HU), and Air Cavity. In addition, the values also consist of a diagnosis that supports the program used, based on 200 observations, which were divided into 2 classes of 102 for acute sinusitis patients and 98 for non-acute. Based on Gender, 0 was stated for male and 1 for female, while on diagnosis, 0 was specified for the patients without acute sinusitis, and 1 for those with it, as shown in the table below:

| Gender | Age | HU   | Air Cavity | Diagnosis |
|--------|-----|------|------------|-----------|
| 1      | 76  | 138  | -1020      | 0         |
| 1      | 76  | 54   | -1022      | 1         |
| 0      | 20  | 38   | -967       | 1         |
| 0      | 20  | 42   | -992       | 1         |
| 0      | 20  | 15   | -987       | 1         |
| 0      | 20  | 23   | -964       | 1         |
| 0      | 20  | 12   | -954       | 1         |
| 0      | 20  | 22   | -890       | 1         |
| 0      | 20  | 24   | -994       | 1         |

2.2 Support Vector Machines
Support Vector Machines (SVM) is a method that is good at prediction, due to the fact that it provides data with high accuracy, through the use of Structural Risk Minimization (SRM) principle. In addition, the idea is based on identification of the best discriminant boundaries of two classes, with a measuring margin, termed hyperplane, given a sample of training dataset $x_i \in \mathbb{R}^N$, while each label is notated $y_i \in \{0,1\}$ for $i = 1,2,\ldots,N$, where N is equivalent to the amount of data. Therefore, a separator is formulated as Equation (1)

$$w \cdot x + b = 0$$ (1)
2.2.1 Support Vector Machines with Soft Margin [24]. Identify the minimum point in Equation (2), and abstract it into Equation (3), creating a formulation called Quadratic Programming (QP) problem.

\[
\min \left( \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} \xi_i \right)
\]  
(2)

With constraint

\[
y_i(w^T x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0 \quad \forall i = 1,2, \ldots, N
\]  
(3)

It is possible to resolve this challenge by various computational techniques, including Lagrange Multiplier which is showed in Equation (4).

\[
\min L(w, b, \xi, \alpha, \beta) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^{N} \xi_i - \sum_{i=1}^{N} \alpha_i [y_i(w^T x_i + b) - 1 + \xi_i] - \sum_{i=1}^{N} \beta_i \xi_i
\]  
(4)

with constraints which are stated in Equation (5), (6), (7), (8), (9), (10), (11).

\[
\alpha_i \geq 0 \quad (5)
\]
\[
\beta_i \geq 0 \quad (6)
\]
\[
\xi_i \geq 0 \quad (7)
\]
\[
1 - \xi_i - y_i(w^T x_i + b) \geq 0 \quad (8)
\]
\[
\alpha_i[1 - \xi_i - y_i(w^T x_i + b)] = 0 \quad (9)
\]
\[
\beta_i \xi_i = 0 \quad (10)
\]
\[
\forall i = 1,2,\ldots, N \quad (11)
\]

The solution of Lagrange Multiplier can be found by searching partial derivative of L to \(w\), \(b\), and \(\xi_i\). Derivate \(L(w, b, \xi, \alpha, \beta)\) respected to \(w\), \(b\), and \(\xi_i\) equal to zero as it is explained in Equation (12), (14), and (16), then the following is obtained in Equation (13), (15), and (17).

\[
\frac{\partial L(w, b, \xi, \alpha, \beta)}{\partial w} = w - \sum_{i=1}^{N} \alpha_i y_i x_i = 0
\]  
(12)

\[
w = \sum_{i=1}^{N} \alpha_i y_i x_i
\]  
(13)

\[
\frac{\partial L(w, b, \xi, \alpha, \beta)}{\partial b} = - \sum_{i=1}^{N} \alpha_i y_i = 0
\]  
(14)

\[
\sum_{i=1}^{N} \alpha_i y_i = 0
\]  
(15)
\[
\frac{\partial L(w, b, \xi, \alpha, \beta)}{\partial \xi_i} = C - \alpha_i - \beta_i = 0
\]  
(16)

\[
\alpha_i = C - \beta_i
\]  
(17)

Substitute them into Equation (4) and it will get a result as in Equation (18)

\[
L(w, b, \xi, \alpha, \beta) = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i, x_j) + \sum_{i=1}^{N} \alpha_i
\]  
(18)

So, the problems above can be written as Equation (19) below:

\[
\max \left( -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i, x_j) + \sum_{i=1}^{N} \alpha_i \right)
\]  
(19)

with constraint in Equation (20) and (21)

\[
\sum_{i=1}^{N} \alpha_i y_i = 0
\]  
(20)

\[
0 \leq \alpha_i \leq C, \forall i = 1, 2, ..., N
\]  
(21)

A kernel function was used to support the method in coping with non-linear separable data. This was, however, defined as

\[
K(x_i, x_j) = \varphi(x_i), \varphi(x_j)
\]

and through its substitution, the following was obtained in Equation (22):

\[
L = \max \left( -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K(x_i, x_j) + \sum_{i=1}^{N} \alpha_i \right)
\]  
(22)

with constraint in Equation (23) and (24)

\[
\sum_{i=1}^{N} \alpha_i y_i = 0
\]  
(23)

\[
0 \leq \alpha_i \leq C, \forall i = 1, 2, ..., N
\]  
(24)

Satisfactory results and outperform conventional classifiers are obtained by a polynomial or RBF kernel, which is stated in [25], both of which are applied in the Equation (25) and (26) below:

\[
K(x_i, x_j) = (t + x_i^T x_j)^d
\]  
(25)

\[
K(x_i, x_j) = \exp \left( -\frac{||x_i - x_j||^2}{2\sigma^2} \right)
\]  
(26)

The solution of the problem is shown in Equation (27) and (28).

\[
f(x_j) = w^* \cdot x_j + b^*
\]  
(27)

\[
f(x_j) = \text{sgn} \left( \sum_{i=1}^{N} \alpha_i^* y_i K(x_i, x_j) + b^* \right)
\]  
(28)

where \(w^*, y_i, \text{and } b^*\) are stated in Equation (29), (30), and (31) below:
$$w^* = \sum_{i=1}^{N} \alpha_i y_i x_i = 0$$  \hspace{1cm} (29)

$$y_i = \sum_{j \in S} \alpha_j^* y_j K(x_i, x_j) + b$$  \hspace{1cm} (30)

$$b^* = \frac{1}{N_s} \sum_{i \in N_s} \left( y_i - \sum_{j \in N_s} \alpha_j^* y_j K(x_i, x_j) \right)$$  \hspace{1cm} (31)

2.3 Fuzzy Support Vector Machine

The Fuzzy variety is an extension of Support Vector Machines method, via a fuzzy membership, which is associated with each training point $x_i$ [26]. Therefore, it is possible to check all the training points in the class treated uniformly, although the effects are different in numerous real applications.

Therefore, Z Rustam has developed a formulation of fuzzy membership, as seen in Equation (32) and (33) [27]:

$$\mu_i A = e^{-\|x\|^2/\sigma^2} \text{ if } x_i \in A$$  \hspace{1cm} (32)

$$\mu_i B = 1 - \mu_i A \text{ if } x_i \in B$$  \hspace{1cm} (33)

where A is a training dataset and B is a testing dataset [27].

The main purpose of Fuzzy Support Vector Machines is doing a preprocessing data. It will produce a new data that is ready to be processed by Support Vector Machines method. Basically, the step in the Fuzzy Support Vector Machines as same as Support Vector Machines, but a data $x_i$ has to be transformed into a new data $x_i'$ which is shown in Equation (34) and (35) below [27]:

$$\text{If } i \in A, \text{ then } x_i' = \mu_i A x_i$$  \hspace{1cm} (34)

$$\text{If } i \in B, \text{ then } x_i' = \mu_i B x_i$$  \hspace{1cm} (35)

3. Experimental Results

The Support Vector Machines and the Fuzzy variety are evaluated with RBF Kernel, and Polynomial, and subsequently a confusion matrix is created.

$T_P$: Number of samples having acute sinusitis diagnosed correctly

$F_P$: Sum of healthy people that were incorrectly identified to have acute sinusitis

$T_N$: Number of healthy individuals correctly spotted

$F_N$: The amount of samples with acute sinusitis that were incorrectly classified as healthy

| Actual Value | Recognize Value |
|--------------|-----------------|
|              | Positive | Negative |
| Positive     | $T_P$     | $F_P$     |
| Negative     | $F_N$     | $T_N$     |

Accuracy explains how a data is being classified, while the second indicates sensitivity, which measures the proportion of actual positive cases that were predicted as such or true. In addition, the
third is precision, which is the positive predictive value, while the last was denoted as F1-Score, used to determine the balance between sensitivity and precision.

The formulas are seen below:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%
\]

\[
\text{Sensitivity} = \text{Recall} = \frac{TP}{TP + FN} \times 100\%
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \times 100\%
\]

\[
\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\%
\]

Based on the results, Table 3 uses ten parameter \( \sigma \), while Table 4 expended ten degrees(\( d \)) as the parameters for each k-fold. These include 1.00E-08, 0.001, 0.05, 0.1, 1, 10, 50, 100, 1000, 10000 for the \( \sigma \) and 1,2,…,10 for the \( d \).

**Table 3.** Optimum classification of Acute Sinusitis Data using Support Vector Machines, and RBF Kernel with parameter \( \sigma \)

| No. | K-fold | \( \sigma \) | Accuracy  | Sensitivity | Precision | F1-Score |
|-----|--------|-------------|-----------|-------------|-----------|----------|
| 1   | 3      | 50          | 96.75     | 96.88       | 96.44     | 96.66    |
| 2   | 5      | 10          | 97.35     | 97.72       | 96.87     | 97.29    |
| 3   | 7      | 10000       | 97.04     | 97.55       | 96.57     | 97.06    |
| 4   | 10     | 1000        | 97.08     | 97.04       | 96.80     | 96.92    |

Based on Table 3, it is recorded that the highest accuracy was seen in \( k = 5, \sigma = 10 \), with 97.35%, while the lowest was observed in \( k = 3, \sigma = 50 \), at 96.75%.

**Table 4.** Optimum classification of Acute Sinusitis Data using Support Vector Machines with Polynomial Kernel and degree \( d \)

| No. | K-fold | \( d \) | Accuracy  | Sensitivity | Precision | F1-Score |
|-----|--------|--------|-----------|-------------|-----------|----------|
| 1   | 3      | 2      | 96.46     | 96.35       | 96.35     | 96.35    |
| 2   | 5      | 1      | 96.92     | 96.84       | 96.84     | 96.84    |
| 3   | 7      | 1      | 97.45     | 97.96       | 96.97     | 97.46    |
| 4   | 10     | 1      | 97.89     | 97.78       | 97.78     | 97.78    |

From Table 4, it was shown that the highest accuracy is in \( k = 10, d = 1 \) at 97.89%, while the lowest was seen in \( k = 3, d = 2 \), with 96.46%. 


Table 5. Optimum classification of Acute Sinusitis Data using Fuzzy Support Vector Machines, and RBF Kernel with parameter $\sigma$

| No. | K-fold | $\sigma$   | Accuracy | Sensitivity | Precision | F1-Score |
|-----|--------|------------|----------|-------------|-----------|----------|
| 1   | 3      | 1.00E-08   | 98.99    | 100         | 97.96     | 98.97    |
| 2   | 5      | 1.00E-08   | 97.95    | 98.95       | 96.91     | 97.92    |
| 3   | 7      | 10000      | 96.94    | 97.96       | 96.00     | 96.97    |
| 4   | 10     | 10000      | 97.16    | 97.78       | 96.28     | 97.02    |

From Table 5, it was recorded that the highest accuracy was observed in $k = 3$, $\sigma = 1.00E - 08$, at 98.99%, while the lowest was seen in $k = 7$, $\sigma = 10000$, at 96.94%.

Table 6. Optimum classification of Acute Sinusitis Data using Fuzzy Support Vector Machines, with Polynomial Kernel and degree $d$

| No. | K-fold | $d$   | Accuracy | Sensitivity | Precision | F1-Score |
|-----|--------|------|----------|-------------|-----------|----------|
| 1   | 3      | 1    | 98.99    | 100         | 97.96     | 98.97    |
| 2   | 5      | 1    | 97.44    | 98.95       | 95.92     | 97.41    |
| 3   | 7      | 1    | 96.94    | 97.96       | 96.00     | 96.97    |
| 4   | 10     | 10   | 97.05    | 97.78       | 96.07     | 96.92    |

From Table 6, it was recorded that the highest accuracy was observed at $k = 3$, $d = 1$ with 98.99%, and the lowest was seen in $k = 7$, $d = 1$, at 96.94%.

The graphs below show a comparison between Accuracy, Sensitivity, Precision, and F1-Score for each method.
Figure 1. Graph of Acute Sinusitis Classification for its Accuracy, Sensitivity, Precision, and F1-Score

Based on Figure 1, the best method at accuracy, sensitivity, precision, and f1-score in Acute Sinusitis Classification are known. In addition, the accuracy, sensitivity, precision, and f1-score in FSVM Polynomial and RBF Kernel were recorded as methods that possessed the highest rate, while that of SVM was the lowest. Hence, based on the entire graph, the advantage observed was due to the fact that it is the extended form, hence, determining each data point with a fuzzy membership.

4. Conclusion
Based on the result and discussion, it is possible to conclude that Fuzzy Support Vector Machines are better at classifying acute sinusitis data, in contrast with the Support Vector Machines, and its usefulness, especially based on accuracy is illustrated with rate. The accuracy of Fuzzy Support Vector Machines with RBF Kernel reach a highest rate at 98.99% while Support Vector Machines with Polynomial Kernel reach a lowest rate at 97.35%. Furthermore, there are expectations that this method enables medical staffs to easily classify a disease or other medical problems. Moreover, subsequent investigations ought to identify other methods, using larger dataset, in order to create models that are more optimal, for the purpose of resolving classification problem.

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

References
[1] Arora A S and Singh J 2017 A Framework for Enhancing the Thermographic Evaluation on Characteristic Areas for Pranasal Sinusitis Detection, Infrared Physics & Technology, 85, 457-464
[2] Abdulaziz A A, Abdulrahman F N A, Adel A A, Ayman S A, Basyuni M A, Muath A A, Rahaf
[3] Arfiani, Pandelaki J, Rustam Z, and Siahaan A 2019 *Kernel Spherical K-Means and Support Vector Machine for Acute Sinusitis Classification*, IOP Conference Series Materials Science and Engineering 546:052011.

[4] Campos C N, de Lima C O, Devita K L, do Prado M, and Vasconcelos L R B 2017 *Correlation between Endodontic Infection and Periodontal Disease and Their Association with Chronic Sinusitis: A Clinical-tomographic Study*, American Association of Endodontists.

[5] Bergmark R W, Campbell A P, and Metson R 2017 *Orbital Complications of Acute Sinusitis*, Operative Techniques in Otolaryngology - Head and Neck Surgery.

[6] Chaudhry Z A, Reede D, Shinder R, Smoker W R K, and Velayudhan V 2017 *Imaging of Intracranial and Orbital Complications of Sinusitis and Atypical Sinus Infection: What the Radiologist Needs to Know*, Current Problems in Diagnostic Radiology.

[7] Lakhan K J T 2018 *Sinus Headaches Sinusitis Versus Migraine*, Physician Assist Clin. 3, 181–192.

[8] Mallon W K and Wyler B 2019 *Sinusitis Update*, Emerg Med Clin N Am, 37, 41–54.

[9] Rachman A A and Rustam Z 2016 *Cancer classification using Fuzzy C-Means with feature selection*, Proceedings, 10.1109/ICMSA.2016.7954302.

[10] Panca V and Rustam Z 2017 *Application of machine learning on brain cancer multiclass classification*, AIP Conference Proceedings, 1862, 10.1063/1.4991237.

[11] Lestari A W and Rustam Z 2017 *Normed kernel function-based fuzzy possibilistic C-means (NKFPCM) algorithm for high-dimensional breast cancer database classification with feature selection is based on Laplacian Score*, AIP Conference Proceedings, 1862, 10.1063/1.4991247.

[12] Acar M, Nadhifa F, and Rustam Z 2018 *Comparison of SVM and FSVM for Predicting Bank Failures using Chi-Square Feature Selection*, Journal of Physics: Conference Series 1108 (1), 012115.

[13] Rukhsan B and Vasile P 2010 *FSVM-CIL Fuzzy Support Vector Machines for Class Imbalance Learning*, IEEE Transactions on Fuzzy Systems, 10.1109/TFUZZ.2010.2042721.

[14] Rustam Z and Zahras D 2018 *Comparison Between Support Vector Machine and Fuzzy C-Means as Classifier for Intrusion Detection System*, Journal of Physics: Conference Series 1028 (1), 012227.

[15] Primasari I, Rustam Z, and Widya D 2018 *Classification of Cancer Data Based on Support Vectors Machines with Feature Selection using Genetic Algorithm and Laplacian Score*, AIP Conference Proceedings, 2023 (1), 020234.

[16] Rustam Z 2018 *Classification of Cancer Data using Support Vector Machines with Feature Selection Method Based on Global Artificial Bee Colony*, AIP Conference Proceedings, 2023 (1), 020205.

[17] Xiao J 2019 *SVM and KNN Ensemble Learning for Traffic Incident Detection*, Physica A 517, 29–35.

[18] Bing G, Chunhui Z, Lejun Z, and Xiaqing W 2018 *Classification of Hyperspectral Imagery Based on Spectral Gradient, SVM and Spatial Random Forest*, Infrared Physics and Technology 95, 61–69.

[19] Rampisela T V and Rustam Z 2018 *Classification of Schizophrenia data using Support Vector Machine (SVM)*, Journal of Physics: Conference Series 1108 (1), 012044.

[20] Faradina R and Rustam Z 2018 *Face Recognition to Identify Look-Alike Faces using Support Vector Machine*, Journal of Physics: Conference Series 1108 (1), 012071.

[21] Rustam Z and Ruvita A A 2018 *Aplication Support Vector Machine on Face Recognition for Gender Classification*, Journal of Physics: Conference Series 1108 (1), 012067.
[22] Rustam Z, Sarwinda D, and Sudarsono E 2019 Random-Forest (RF) and Support Vector Machine (SVM) Implementation for Analysis of Gene Expression Data in Chronic Kidney Disease (CKD), IOP Conference Series Materials Science and Engineering 546:052066

[23] Rustam Z and Yaurita F 2018 Insolvency Prediction in Insurance Companies Using Support Vector Machines and Fuzzy Kernel C-Means, Journal of Physics: Conference Series 1028 (1), 012118

[24] Dea A U, 2019. Klasifikasi Imbalanced Data Infark Serebri Menggunakan Support Vector Machine. Thesis

[25] Hofmann M 2006 Support vector machines–kernels and kernel trick Houptseminar Report

[26] Chi-Fan L and Sheng-De W 2002 Fuzzy Support Vector Machine, IEEE Transaction on Neural Networks, 10.1109/72.991432

[27] Farah N, 2018. Klasifikasi Future Claim pada Asuransi Kendaraan Bermotor dengan Menggunakan Metode Support Vector Machines dan Fuzzy Support Vector Machines. Thesis