Kernel perceptron algorithm for sinusitis classification

Z Rustam¹, S Hartini¹, and J Pandelaki²
¹Department of Mathematics, University of Indonesia, Depok 16424, Indonesia
²Department of Radiology, Cipto Mangunkusumo Hospital, Jakarta 10430, Indonesia

rustam@ui.ac.id

Abstract. Sinusitis is one of the most commonly diagnosed diseases in the world. Its diagnosis is usually based on clinical signs and symptoms, which led to the development and use of many machine learning methods to provide a better diagnosis. This research, therefore, proposed a kernel perceptron method applied to the sinusitis dataset, consisting of 102 acute and 98 chronic samples, obtained from Cipto Mangunkusumo Hospital in Indonesia. This research utilized the RBF and polynomial kernel function for several k values in k-fold cross-validation and compared the results in accuracy, sensitivity, precision, specificity, and F1-Score. From the experiments, it was concluded that the kernel parameter \( \sigma = 0.0001 \) obtained excellent performance in every k-fold, with a better performance achieved using 10-fold cross-validation. Meanwhile, the polynomial degree did not affect the kernel perceptron performance. However, the use of 7-fold cross-validation can be considered to obtain better performance of kernel perceptron based on polynomial kernel.

1. Introduction

Sinusitis, also known as rhinosinusitis, is a nasal inflammation. According to Ah-See et al. [1], the European Academy of Allergology and Clinical Immunology defines acute rhinosinusitis as the inflammation of the nose and the paranasal sinuses which lasts less than 12 weeks and accompanied by two or more symptoms such as blockage/congestion, discharge, facial pain, reduction or loss of smell. Chronic sinusitis is nasal congestion characterized by facial pain, discolored nasal discharge, and reduction or loss of smell for more than 12 weeks.

Sinusitis is believed to affect more than 12% of the US population [2], while the prevalence of chronic sinusitis is estimated at 2.1% in the eight major cities in China and 8% in seven others [3]. However, it affects approximately 5–15% of the general population [4]. In most cases, patients recover without antibiotic therapy [5]. Therefore, treatment should be reserved for those with symptoms for at least 7-10 days and treated with conservatively, while those with chronic sinusitis are treated differently [6].

This ailment was classified into kernel spherical k-means and support vector machine [7], convolutional neural network [8,9], and fuzzy kernel c-means [10] by researchers. The kernel function has already been used in many papers and applied to several machine learning methods [11-15]. Therefore, this study proposes the kernel perceptron for sinusitis classification by modifying the perceptron algorithm using the kernel function.
2. Material and Methods

2.1. Material

Sinusitis dataset used consists of 102 acute and 98 chronic sinusitis samples, obtained from Cipto Mangunkusumo (RSCM) Hospital, Indonesia. The instances were described by four features: gender, age, Hounsfield unit, and air cavity.

2.2. Methods

2.2.1. Kernel Perceptron. Perceptron learning algorithm was first introduced by Aizerman et al. [16] in 1964. However, Ortiz et al. [17] applied the kernel function to it and named the new method as a kernel perceptron. After applying the kernel function, it allows learning machines to classify data that cannot separate linearly. The kernel function for every \( x \in \mathbb{R}^n \) is given below [18].

\[
K(x, x') = \langle \phi(x), \phi(x') \rangle
\]

where \( \phi \) is the mapping function from the input data samples to the feature space. Although there are several kernel functions, this research utilized the RBF and polynomial kernel function with the formulas defined in Eq. (2) and Eq. (3), respectively. After that, the performance of each kernel function in the kernel perceptron is compared.

RBF kernel function:

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

Polynomial kernel function:

\[
K(x_i, v_j) = (x_i \cdot v_j + 1)^h
\]

The classification function is given as follow:

\[
f(x) = \text{sign} \left( \sum_{i=1}^{l} \alpha_i y_i x_i \cdot x + b \right)
\]

or

\[
f(x) = \text{sign}(K(w, x) + b)
\]

where \( \alpha_i \) is the Lagrange multiplier for the \( i \)-th sample, and the algorithm of kernel perceptron [19] is shown in Figure 1.

Input: Data training \( X = \{ x_1, x_2, \ldots, x_m \} \), \( x_i \in \mathbb{R}^d \), the label of the \( i \)-th sample \( y_i \in [-1, 1] \), and learning rate \( \eta \in \mathbb{R}^+ \)

Output: \( w \in \mathbb{R}^d \) and \( b \in \mathbb{R} \)

1. Initialization: \( w_0 = 0, b_0 = 0 \), OK = TRUE
2. Set \( R = \max_{i \in [1]} \| x_i \| \)
3. WHILE OK DO
   4. FOR \( i = 1 \) : \( m \)
      5. IF \( f(x_i) \cdot y_i \leq 0 \), for all samples in the dataset, update \( w \) and \( b \) using:
         6. \( w_{k+1} = w_k + \eta y_i x_i \)
         7. \( b_{k+1} = b_k + \eta y_i R^2 \)
         8. OK = TRUE
   9. ELSE
      10. OK = FALSE
11. END IF
12. END FOR
13. END WHILE
14. End

Figure 1. Algorithm of kernel perceptron
2.2.2. \textit{k-fold cross-validation}. The evaluation of the performances of kernel perceptron using RBF and polynomial kernel function is committed using k-fold cross-validation where \( k = 3, 5, 7, 10 \). Using k-fold cross-validation, the dataset was divided randomly into \( k \) folds with an approximately equal number of samples [20]. In every iteration, each fold was taken as the validation data for testing the kernel perceptron model, and the rest of the dataset was chosen as the training data.

2.2.3. \textit{Performance measure}. The measurement of kernel perceptron performance in sinusitis classification is defined in Eqs. (6)-(10).

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FN + FP} \quad (6)
\]
\[
\text{Sensitivity} = \frac{TP}{TP + FN} \quad (7)
\]
\[
\text{Precision} = \frac{TP}{TP + FP} \quad (8)
\]
\[
\text{Specificity} = \frac{TN}{TN + FP} \quad (9)
\]
\[
\text{F1-Score} = \frac{2 \times \text{Sensitivity} \times \text{Precision}}{\text{Sensitivity} + \text{Precision}} \quad (10)
\]

where TP is the number of chronic sinusitis samples which were correctly diagnosed, FP is the total of acute sinusitis samples which were incorrectly diagnosed, TN is the count of acute sinusitis samples which were correctly diagnosed, and FN is the number of chronic sinusitis samples which were incorrectly diagnosed.

Consider the relevant case in this paper as chronic sinusitis. Kotu and Deshpande [21] described each of the performance measurement abilities as follows. Accuracy is the aggregate measure of classifier performance, while sensitivity is used to select what requires to be selected, and specificity used to reject what needs to be rejected. Furthermore, they defined precision as the proportion of cases found that was relevant and the recall as the proportion of all relevant cases that were found.

3. \textbf{Results and Discussions}

The \( k \)-fold cross-validation where \( k = 3, 5, 7, \) and 10 were used to evaluate the performance of kernel perceptron, were all based on RBF and polynomial kernel function. The performance of the kernel perceptron using RBF kernel function is shown in Table 1.

\textbf{Table 1. The performance of kernel perceptron using RBF kernel function}

| Evaluation method | Kernel Parameter | Performance measure | Running time |
|-------------------|------------------|---------------------|--------------|
|                   |                  | Accuracy  | Sensitivity | Precision | Specificity | F1-Score |             |
| 3-fold CV         | 0.0001           | 96.97     | 97.92      | 95.92     | 96.08      | 96.91    | 1.13        |
|                   | 0.001            | 96.72     | 96.88      | 96.37     | 96.57      | 96.62    | 0.03        |
|                   | 0.05             | 96.63     | 96.88      | 96.21     | 96.41      | 96.54    | 0.09        |
|                   | 0.1              | 96.59     | 96.88      | 96.12     | 96.32      | 96.50    | 0.05        |
|                   | 1                | 96.57     | 96.88      | 96.07     | 96.27      | 96.47    | 0.03        |
|                   | 5                | 96.55     | 96.88      | 96.04     | 96.24      | 96.46    | 0.03        |
|                   | 10               | 96.54     | 96.88      | 96.02     | 96.22      | 96.44    | 0.03        |
|                   | 50               | 96.53     | 96.88      | 96.00     | 96.20      | 96.44    | 0.02        |
|                   | 100              | 96.52     | 96.88      | 95.99     | 96.19      | 96.43    | 0.03        |
|                   | 1000             | 96.52     | 96.88      | 95.98     | 96.18      | 96.42    | 0.05        |
| 5-fold CV         | 0.0001           | 96.92     | 96.84      | 96.84     | 97.00      | 96.84    | 1.34        |
|                   | 0.001            | 95.90     | 94.74      | 96.77     | 97.00      | 95.74    | 0.09        |
|                   | 0.05             | 96.24     | 95.44      | 96.80     | 97.00      | 96.11    | 0.06        |
|                   | 0.1              | 96.41     | 95.79      | 96.81     | 97.00      | 96.30    | 0.03        |
|                   | 1                | 96.51     | 96.00      | 96.82     | 97.00      | 96.41    | 0.05        |
|                   | 5                | 96.58     | 96.14      | 96.82     | 97.00      | 96.48    | 0.03        |
|                   | 10               | 96.63     | 96.24      | 96.82     | 97.00      | 96.53    | 0.05        |
|                   | 50               | 96.67     | 96.32      | 96.83     | 97.00      | 96.57    | 0.03        |
According to this table, the kernel parameter $\sigma = 0.0001$ reaches the best performance in every type of evaluation. It achieves the accuracy, sensitivity, precision, specificity, and F1-score above 95 percent. Therefore, kernel perceptron based on RBF kernel function with $\sigma = 0.0001$ is the best performance for all tested values. The comparison of kernel perceptron for each type of evaluation is given in Figure 2.

| $k$-fold CV  | 3-fold CV   | 5-fold CV   | 7-fold CV   | 10-fold CV  |
|--------------|-------------|-------------|-------------|-------------|
|              | 96.97       | 97.92       | 95.92       | 96.08       |
| 0.0001       | 96.94       | 96.94       | 96.94       | 96.94       |
| 0.001        | 95.92       | 94.90       | 96.88       | 96.94       |
| 0.05         | 96.26       | 95.58       | 96.90       | 96.94       |
| 0.1          | 96.43       | 95.92       | 96.91       | 96.94       |
| 1            | 96.53       | 96.12       | 96.91       | 96.94       |
| 5            | 96.60       | 96.26       | 96.92       | 96.94       |
| 10           | 96.65       | 96.36       | 96.92       | 96.94       |
| 50           | 96.68       | 96.43       | 96.92       | 96.94       |
| 100          | 96.71       | 96.49       | 96.92       | 96.94       |
| 1000         | 96.73       | 96.53       | 96.93       | 96.94       |

According to this table, the kernel parameter $\sigma = 0.0001$ reaches the best performance in every type of evaluation. It achieves the accuracy, sensitivity, precision, specificity, and F1-score above 95 percent. Therefore, kernel perceptron based on RBF kernel function with $\sigma = 0.0001$ is the best performance for all tested values. The comparison of kernel perceptron for each type of evaluation is given in Figure 2.

![Figure 2. Comparison of k-fold cross-validation of kernel perceptron using RBF kernel function with $\sigma = 0.0001$](image-url)

In this figure, the value of the performance measurement grows as the $k$ value in k-fold cross-validation increases for several cases, such as accuracy and F1-Score. Moreover, the best performance measurement is achieved for different values of $k$ in the k-fold cross-validation. However, the best value of accuracy, specificity, and F1-Score is obtained when we evaluated with 10-fold cross-validation. The performance of the kernel perceptron based on polynomial kernel function is shown in Table 2.
According to Table 2, kernel perceptron based on polynomial kernel function is not affected by the polynomial degree used. However, the comparison of kernel perceptron based on polynomial kernel for every k-fold cross-validation is shown in Figure 3.

| Evaluation method | Polynomial degree | Performance measure | Running time |
|-------------------|-------------------|----------------------|--------------|
|                   |                   | Accuracy | Sensitivity | Precision | Specificity | F1-Score | |
| 3-fold CV         | 1                 | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.05       |
|                   | 2                 | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.03       |
|                   | 3                 | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.16       |
|                   | 4                 | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.03       |
|                   | 5                 | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.03       |
|                   | 6                 | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.05       |
|                   | 7                 | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.03       |
|                   | 8                 | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.06       |
|                   | 9                 | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.06       |
|                   | 10                | 96.46    | 96.88      | 95.88      | 96.08      | 96.37    | 0.03       |
| 5-fold CV         | 1                 | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.08       |
|                   | 2                 | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.03       |
|                   | 3                 | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.03       |
|                   | 4                 | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.03       |
|                   | 5                 | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.03       |
|                   | 6                 | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.03       |
|                   | 7                 | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.03       |
|                   | 8                 | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.05       |
|                   | 9                 | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.05       |
|                   | 10                | 96.92    | 96.84      | 96.84      | 97.00      | 96.84    | 0.03       |
| 7-fold CV         | 1                 | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.09       |
|                   | 2                 | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.06       |
|                   | 3                 | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.06       |
|                   | 4                 | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.06       |
|                   | 5                 | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.05       |
|                   | 6                 | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.05       |
|                   | 7                 | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.05       |
|                   | 8                 | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.03       |
|                   | 9                 | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.03       |
|                   | 10                | 96.94    | 96.94      | 96.94      | 96.94      | 96.94    | 0.03       |
| 10-fold CV        | 1                 | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.08       |
|                   | 2                 | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.03       |
|                   | 3                 | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.06       |
|                   | 4                 | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.03       |
|                   | 5                 | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.03       |
|                   | 6                 | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.03       |
|                   | 7                 | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.06       |
|                   | 8                 | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.03       |
|                   | 9                 | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.05       |
|                   | 10                | 96.84    | 96.67      | 96.67      | 97.00      | 96.67    | 0.08       |
Figure 3. Comparison of k-fold cross-validation of kernel perceptron using the polynomial kernel function

Kernel perceptron using polynomial kernel function performs excellently for every k-fold cross-validation. However, according to the figure above, the 7-fold cross-validation makes kernel perceptron based on polynomial kernel performs better than the others in the accuracy, sensitivity, precision, and F1-Score.

4. Conclusion
Sinusitis is a nasal inflammation, which can be classified as acute and chronic sinusitis, affects many people in this world. Various machine learning methods have been developed to deliver a better result in its classification. However, the possibility to compare the performance of the machine learning method is always open. Therefore, this research proposed a kernel perceptron method using a sinusitis dataset consisting of 102 acute and 98 chronic samples obtained from Cipto Mangunkusumo Hospital in Indonesia.

Several k values in k-fold cross-validation for RBF and polynomial kernel function in the kernel perceptron algorithm was utilized. From the experiments, it is concluded that the kernel parameter $\sigma = 0.0001$ reaches excellent performance in every k-fold. However, the performance is better using 10-fold cross-validation. Meanwhile, no effect was found on the polynomial degree used in polynomial kernel function. However, the use of 7-fold cross-validation can be considered to makes kernel perceptron based on polynomial kernel performs better.

Other methods or kernel functions can be used in future experiments. The evaluation method can also be modified to understand how this classification method works better. Overall, the more accurate and precise diagnosis is expected to obtain better work in the health field.

Acknowledgment
The authors were grateful to the Department of Radiology, Cipto Mangunkusumo Hospital for their kindness in providing the sinusitis datasets, and all reviewers included in the improvement of this article. The authors were also thankful for the financial support from the Ministry of Research, Technology, and Higher Education Republic of Indonesia, with the PDUPT 2019 research grant scheme (ID Number 1621/UN2.R3.1/HKP.05.00/2019).
References

[1] Ah-See K W and Evans A S 2007 BMJ 334 358-361
[2] Shashy R G, Moore E J, and Weaver A 2004 Otolaryng. Head Neck Surg. 130 320-323
[3] Zhang Y, Gevaert E, Lou H, Wang X, Zhang L, Bachert C, and Zhang N 2017 J. Allergy Clin. Immunol. 140 1230-1239
[4] Bachert et al. 2014 World Allergy Organ. J. 7 1-28
[5] Wilson JF 2010 Ann. Intern. Med. 153 ITC3-1-15
[6] Mustafa M, Patawari P, Ifikhar HM, Shimmi SC, Hussain SS, and Sien MM 2015 Int. J. Pharm. Sci. Drug Res. 4 30-36
[7] Arfiani, Rustam Z, Pandelaki J, and Sihaan A 2019 IOP Conf. Ser.: Mater. Sci. Eng. 546 052011
[8] Kim H-G, Lee K M, Kim E J, and Lee J S 2019 Quant Imaging Med Surg 9 942-951
[9] Chowdhury N I, Smith T L, Chandra R K, and Turner J H 2019 Int. Forum Allergy Rhinol. 9 46–52
[10] Putri R A, Rustam Z, and Pandelaki J 2019 IOP Conf. Ser.: Mater. Sci. Eng. 546 052060
[11] Rustam Z and Talita A S 2018 3rd Int. Conf. on Inf. and Comput. pp 1-4
[12] Shandri N and Rustam Z 2018 Int. Conf. on Appl. Inf. Technol. and Innov. pp 145-148
[13] Rustam Z and Hartini S 2019 IOP Conf. Ser.: Mater. Sci. Eng. 546 052067
[14] Wulan A, Jannati M V, Rustam Z, and Fauzan A A 2016 Proc. Int. Conf. on Math. Stat. and Their App. pp 35-38
[15] Rustam Z and Olivera N 2018 AIP Conf. Proc. 2023 020215
[16] Aizerman M A 1964 Automat. Rem. Contr. 25 pp 821-837
[17] Ortiz N, Jimenez-Moreno R, and Mauleodux M 2018 Int. J. Appl. Eng. 13 14698-14700
[18] Vapnik V N 1998 Statistical Learning Theory (John Wiley & Sons, Hoboken, New Jersey)
[19] Cristianini N and Shawe-Taylor J 2014 An introduction to support vector machines and other kernel-based learning methods (Cambridge: Cambridge University Press) pp 26-51
[20] Kohavi R 1995 Proc. 14th Int. Joint Conf.on Artif. Intell. (Montreal) 2 pp 1137-1143
[21] Kotu V and Deshpande B 2019 Model Evaluation Data Science (Second Edition) eds Kotu V and Deshpande B (Cambridge: Morgan Kaufmann) chapter 8 pp 263-279