Schizophrenia Classification Using Fuzzy Kernel C-Means

S Hartini* and Z Rustam

1Department of Mathematics, University of Indonesia, Depok 16424, Indonesia

*e-mail: sri.hartini@sci.ui.ac.id

Abstract. Schizophrenia is a severe mental disorder that induces the mind, feeling, and behavior. Earlier treatment for this disease is essential; therefore, the ability to predict schizophrenia also becomes essential. Fuzzy kernel c-means was proposed in this research, utilizing the data obtained from Northwestern University, which consists of 171 schizophrenia and 221 non-schizophrenia samples. There are several kernel functions; however, RBF and polynomial kernel function were used. As the evaluation, k-fold cross-validation with k = 3, 5, 7, and 10 was used, and each of the performances was analyzed. From the experiments, it was concluded that fuzzy kernel c-means using RBF kernel with σ = 0.01 and σ = 1 provides better performance than polynomial kernel with a similar running time with fuzzy c-means.

Keyword: Schizophrenia classification, fuzzy kernel C-means

1. Introduction

Schizophrenia attacks the mind, feeling, and behavior of someone and distract the ability to be involved in social life. It is a severe mental disorder that affects more than 21 million people worldwide, but more than 50% of people with schizophrenia are not receiving appropriate care [1]. However, the treatment for someone that has schizophrenia is essential. Wu et al. [2] assumed that a focus on improving controllable factors, including medication adherence and physical health status, might help to prevent aggressive behavior in a patient with schizophrenia.

Several machine learning method has used to distinguish between schizophrenia patients with healthy people. Support Vector Machines (SVM) and the random forest [3] were given an excellent performance that above 70 percent accuracy. After that, 80 percent accuracy can be obtained using multimodal markers of schizophrenia, machine learning techniques that decision-based [4]. In another study, Magnetic Resonance Imaging (MRI) brain scans dataset was applied on various supervised machine learning classifiers [5], and the results show more than 90 percent accuracy. Another SVM [6] and Twin SVM [7] also used in different schizophrenia datasets and were given more than 90 percent accuracy.

It seems that the common classification methods have already delivered an excellent performance in distinguishing schizophrenia and healthy sample. Therefore, in this research, instead of using the classification method, the fuzzy kernel c-means clustering was proposed and expected to give a more accurate diagnosis.
2. Material and Methods

2.1. Material

The Northwestern University Schizophrenia dataset [8] was used in this research. This dataset consists of 171 schizophrenia and 221 non-schizophrenia samples, which are described by 64 features. The features are 34 answers of the Scale for the Assessment of Positive Symptoms (SAPS), 25 answers of the Scale for the Assessment of Negative Symptoms (SANS), gender, dominant hand, race, ethnic, and age of the patients.

2.2. Methods

2.2.1. Fuzzy c-means. The fuzzy c-means clustering algorithm is the most commonly used method in the clustering field. This method initially been due to Dunn [9] and later modified by Bezdek [10]. This method divides the dataset into \( c \) clusters. Given \( n \) samples from a dataset \( \{x_1, x_2, ..., x_n\} \) and \( c \) centers of the clusters \( \{v_1, v_2, ..., v_c\} \). Suppose \( u_{ij} \epsilon [0,1] \) is a membership of the data point \( x_i \) in the \( j \)-th cluster that satisfies Eqs. (1)-(2) where \( c < n \) is a positive integer.

\[
\sum_{j=1}^{c} u_{ij} = 1, \quad i = 1, 2, ..., n \quad (1)
\]
\[
0 < \sum_{i=1}^{n} u_{ij} < n, \quad j = 1, 2, ..., c \quad (2)
\]

Then the objective function of fuzzy c-means was given below.

\[
J_m = \sum_{i=1}^{n} \sum_{j=1}^{c} (u_{ij})^m d^2(x_i, v_j) \quad (3)
\]

The algorithm of this method is shown in Figure 1.

```
Input: the dataset \( X = \{x_1, x_2, ..., x_n\} \), number of clusters \( c \), the maximum number of iterations \( T \), fuzziness degree \( m \), cutoff threshold \( 0 < \epsilon < 1 \)
Output: \( V = \{v_1, v_2, ..., v_c\} \), \( U = [u_{ij}] \), \( 1 \leq i \leq n, 1 \leq j \leq c \).
1. Initialization: \( V^0 = \{v_1, v_2, ..., v_c\} \)
2. Update membership of the data point \( x_i \) in \( j \)-th cluster
\[
u_{ij} = \left( \frac{d(x_i, v_j)}{\sum_{j=1}^{c} \left( \frac{d(x_i, v_j)}{m-1} \right)^{-\frac{2}{m-1}}} \right)^{-\frac{2}{m-1}}, \quad 1 \leq i \leq n, 1 \leq j \leq c
\]
3. Update cluster center \( V^t \)
\[
v_j = \frac{\sum_{i=1}^{n} u_{ij}^m x_i}{\sum_{i=1}^{n} u_{ij}^m}, \quad j = 1, 2, ..., c
\]
4. If \( \|V^{(t-1)} - V^{(t)}\| < \epsilon \) or \( T = t \), then the iteration stops. Otherwise, \( t = t + 1 \) and go back to step 2;
5. End.
```

**Figure 1.** The algorithm of fuzzy c-means

2.2.2. Fuzzy kernel c-means. Fuzzy kernel c-means algorithm was built by applying the kernel-based distance into fuzzy c-means [11]. It has the same concept as fuzzy c-means clustering with the objective function that was defined as follows.
In this study, two kernel functions were used, namely the polynomial kernel function and Gaussian Radian Basis Function (RBF). The polynomial kernel function with the first ten polynomial degree \( h \) were examined as same as the sigma value in the RBF kernel function. The polynomial and RBF kernel function formulas are given in Eqs. (5)-(6), respectively.

\[
K(x, y) = (x \cdot y + 1)^h \\
K(x, y) = \exp \left( -\frac{\|x-y\|^2}{2\sigma^2} \right)
\]

The algorithm of fuzzy kernel c-means is given in Figure 2.

1. Initialization: \( V^0 = \{v_1, v_2, \ldots, v_c\} \)
2. Update membership of the data point \( x_i \) in \( j \)th-cluster
   \[
   u_{ij} = \frac{(K(x_i,x_j) - 2K(x_i,v_j) + K(v_j,v_j))^{-\frac{1}{m-1}}}{\sum_{j=1}^{c}(K(x_i,x_j) - 2K(x_i,v_j) + K(v_j,v_j))^{-\frac{1}{m-1}}} , 1 \leq i \leq n, 1 \leq j \leq c
   \]
3. Update cluster center \( V^t \)
   \[
   v_j = \frac{\sum_{i=1}^{n} u_{ij}^m x_i}{\sum_{i=1}^{n} u_{ij}^m} , j = 1, 2, \ldots, c
   \]
4. If \( \|V^{(t-1)} - V^{(0)}\| < \varepsilon \) or \( T = t \), then the iteration stops. Otherwise, \( t = t + 1 \) and go back to step 2;
5. End.

**Figure 2.** The algorithm of fuzzy kernel c-means

**2.2.3. Performance measure.** In order to evaluate the performance of fuzzy c-means and fuzzy kernel c-means, the confusion matrix [12] was used. Because the schizophrenia dataset was used in this research, therefore, the confusion matrix used was given in Figure 3.

| Actual Class | Schizophrenia | Non-Schizophrenia |
|--------------|---------------|-------------------|
| Schizophrenia| TP (True Positive) | FP (False Positive) |
| Non-Schizophrenia | FN (False Negative) | TN (True Negative) |

**Figure 3.** The confusion matrix related to schizophrenia dataset

The information about the number of TP, TN, FP, and FN from the confusion matrix then used to calculate the accuracy, sensitivity, precision, and F1-Score using the formulas in Eqs. (7)-(10).

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \\
\text{Sensitivity} = \frac{TP}{TP + FN} \\
\text{Precision} = \frac{TP}{TP + FP} \\
\text{F1-Score} = \frac{2 \times \text{sensitivity} \times \text{precision}}{\text{sensitivity} + \text{precision}}
\]
3. Results and Discussions

The performance of fuzzy c-means and fuzzy kernel c-means was compared in accuracy, sensitivity, precision, F1-Score, and running time after \( k \)-fold cross-validation was applied to each method. The performance of fuzzy c-means on the schizophrenia dataset was given in Table 1.

| \( k \) | Accuracy | Sensitivity | Precision | F1-Score | Running Time |
|-------|----------|-------------|-----------|----------|--------------|
| 3     | 61.98    | 69.83       | 84.67     | 76.52    | 0.05         |
| 5     | 61.49    | 69.27       | 84.54     | 76.10    | 0.05         |
| 7     | 61.73    | 69.54       | 84.59     | 76.29    | 0.05         |
| 10    | 69.81    | 78.70       | 85.99     | 81.78    | 0.06         |

From this table, 10-fold cross-validation delivers better performance than the other \( k \) values in \( k \)-fold cross-validation. After that, the performance of fuzzy kernel c-means with a polynomial kernel function is shown in Table 2. From this table, the best performer in every \( k \)-fold cross-validation in fuzzy kernel c-means is when using the tenth polynomial kernel function.

| \( k \) | Polynomial Degree | Accuracy | Sensitivity | Precision | F1-Score | Running Time |
|-------|-------------------|----------|-------------|-----------|----------|--------------|
| 3     | 1                 | 61.98    | 69.83       | 84.66     | 76.50    | 0.34         |
| 5     | 2                 | 64.03    | 72.13       | 85.07     | 78.03    | 0.22         |
| 7     | 3                 | 68.88    | 77.59       | 85.96     | 81.52    | 1.25         |
| 10    | 4                 | 72.19    | 81.32       | 86.54     | 83.83    | 1.73         |
|       | 5                 | 76.28    | 85.92       | 87.14     | 86.47    | 1.57         |
|       | 6                 | 80.10    | 90.23       | 87.70     | 88.94    | 1.72         |
|       | 7                 | 80.87    | 91.09       | 87.81     | 89.41    | 1.12         |
|       | 8                 | 81.89    | 92.24       | 87.94     | 90.02    | 1.26         |
|       | 9                 | 82.91    | 93.39       | 88.07     | 90.63    | 1.01         |
|       | 10                | 83.42    | 93.97       | 88.13     | 90.93    | 1.51         |
| 5     | 1                 | 62.25    | 70.12       | 84.71     | 76.71    | 0.21         |
| 7     | 2                 | 63.78    | 71.84       | 85.02     | 77.85    | 0.13         |
|       | 3                 | 69.14    | 77.88       | 86.01     | 81.72    | 1.50         |
|       | 4                 | 73.47    | 82.76       | 86.74     | 84.69    | 1.70         |
|       | 5                 | 77.05    | 86.79       | 87.27     | 86.99    | 1.78         |
|       | 6                 | 80.36    | 90.52       | 87.73     | 89.05    | 1.93         |
|       | 7                 | 80.81    | 90.81       | 87.76     | 89.22    | 1.70         |
|       | 8                 | 81.89    | 92.24       | 87.93     | 90.00    | 1.66         |
|       | 9                 | 82.91    | 93.40       | 88.06     | 90.62    | 1.61         |
|       | 10                | 83.42    | 93.98       | 88.13     | 90.92    | 1.59         |
| 7     | 1                 | 62.24    | 70.14       | 84.71     | 76.66    | 0.19         |
| 10    | 2                 | 63.78    | 71.87       | 85.01     | 77.79    | 0.17         |
|       | 3                 | 69.64    | 78.47       | 86.12     | 82.06    | 1.96         |
|       | 4                 | 72.70    | 81.91       | 86.62     | 84.15    | 1.72         |
|       | 5                 | 77.30    | 87.08       | 87.31     | 87.13    | 2.19         |
|       | 6                 | 80.36    | 90.52       | 87.73     | 89.08    | 1.61         |
|       | 7                 | 80.81    | 90.80       | 87.77     | 89.24    | 1.61         |
|       | 8                 | 82.14    | 92.52       | 87.97     | 90.18    | 1.44         |
|       | 9                 | 82.65    | 93.10       | 88.03     | 90.49    | 1.76         |
|       | 10                | 83.67    | 94.25       | 88.16     | 91.09    | 1.70         |
| 10    | 1                 | 62.27    | 70.15       | 84.62     | 76.52    | 0.19         |
|       | 2                 | 64.06    | 72.17       | 85.00     | 77.88    | 0.29         |
|       | 3                 | 69.17    | 77.92       | 85.97     | 81.60    | 2.75         |
The RBF kernel function used in fuzzy kernel c-means was also examined with only several kernel parameter sigma. According to table 3, the best performer in every \( k \)-fold cross-validation more varied than when using the polynomial kernel function. For several values of sigma, the percentage of accuracy, sensitivity, precision, and F1-Score are constant even though the value of sigma changes. Therefore, the best performer was taken based on the fastest running time.

### Table 3. The performance of fuzzy kernel c-means using polynomial kernel function on schizophrenia dataset

| \( k \) | Sigma | Accuracy | Sensitivity | Precision | F1-Score | Running Time |
|-------|-------|----------|-------------|-----------|----------|--------------|
| 3     | 0.0001| 88.78    | 100.00      | 88.78     | 94.05    | 0.05         |
|       | 0.001 | 88.78    | 100.00      | 88.78     | 94.05    | 0.03         |
|       | 0.05  | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 0.01  | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 1     | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 5     | 56.88    | 64.08       | 83.45     | 72.40    | 0.04         |
|       | 10    | 60.97    | 68.68       | 84.41     | 75.67    | 0.02         |
|       | 50    | 61.99    | 69.83       | 84.63     | 76.46    | 0.02         |
|       | 100   | 61.99    | 69.83       | 84.63     | 76.46    | 0.02         |
|       | 1000  | 57.39    | 64.66       | 83.57     | 72.81    | 0.01         |
| 5     | 0.0001| 88.78    | 100.00      | 88.78     | 94.05    | 0.13         |
|       | 0.001 | 88.78    | 100.00      | 88.78     | 94.05    | 0.03         |
|       | 0.05  | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 0.01  | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 1     | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 5     | 57.39    | 64.65       | 83.62     | 72.89    | 0.03         |
|       | 10    | 60.72    | 68.40       | 84.39     | 75.54    | 0.02         |
|       | 50    | 62.25    | 70.12       | 84.71     | 76.71    | 0.02         |
|       | 100   | 62.25    | 70.12       | 84.71     | 76.71    | 0.02         |
|       | 1000  | 57.14    | 64.36       | 83.56     | 72.69    | 0.01         |
| 7     | 0.0001| 88.78    | 100.00      | 88.78     | 94.05    | 0.09         |
|       | 0.001 | 88.78    | 100.00      | 88.78     | 94.05    | 0.03         |
|       | 0.05  | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 0.01  | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 1     | 77.55    | 85.71       | 88.69     | 94.00    | 0.02         |
|       | 5     | 57.14    | 64.34       | 83.46     | 72.59    | 0.03         |
|       | 10    | 60.97    | 68.65       | 84.32     | 75.58    | 0.02         |
|       | 50    | 62.24    | 70.08       | 84.59     | 76.57    | 0.02         |
|       | 100   | 62.24    | 70.08       | 84.59     | 76.57    | 0.02         |
|       | 1000  | 57.40    | 64.63       | 83.53     | 72.80    | 0.01         |
| 10    | 0.0001| 88.78    | 100.00      | 88.78     | 94.05    | 0.12         |
|       | 0.001 | 88.78    | 100.00      | 88.78     | 94.05    | 0.04         |
|       | 0.05  | 88.78    | 100.00      | 88.78     | 94.05    | 0.04         |
|       | 0.01  | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 1     | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
|       | 5     | 65.79    | 74.13       | 85.14     | 78.75    | 0.06         |
|       | 10    | 61.00    | 68.71       | 84.34     | 75.56    | 0.04         |
|       | 50    | 62.27    | 70.15       | 84.62     | 76.52    | 0.03         |
|       | 100   | 62.27    | 70.15       | 84.62     | 76.52    | 0.02         |
|       | 1000  | 65.53    | 73.83       | 85.09     | 78.57    | 0.02         |
After getting the best performer for all the methods, then the value of them was evaluated and compared. This comparison was given in Table 4. From this table, we can conclude that using RBF kernel function in fuzzy kernel c-means was better than fuzzy c-means and fuzzy kernel c-means using polynomial kernel function.

Table 4. The performance comparison between fuzzy c-means and the best performer from each of fuzzy kernel c-means using polynomial and RBF kernel function on schizophrenia dataset

| k  | The distance used in FCM                                      | Accuracy | Sensitivity | Precision | F1-Score | Running Time |
|----|---------------------------------------------------------------|----------|-------------|-----------|----------|--------------|
| 3  | Euclidean                                                    | 61.98    | 69.83       | 84.67     | 76.52    | 0.05         |
|    | 10th polynomial kernel                                       | 83.42    | 93.97       | 88.13     | 90.93    | 1.51         |
|    | RBF kernel with $\sigma = 0.01$ and $\sigma = 1$            | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
| 5  | Euclidean                                                    | 61.49    | 69.27       | 84.54     | 76.10    | 0.05         |
|    | 10th polynomial kernel                                       | 83.42    | 93.98       | 88.13     | 90.92    | 1.59         |
|    | RBF kernel with $\sigma = 0.01$ and $\sigma = 1$            | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
| 7  | Euclidean                                                    | 61.73    | 69.54       | 84.59     | 76.29    | 0.05         |
|    | 10th polynomial kernel                                       | 83.67    | 94.25       | 88.16     | 91.09    | 1.70         |
|    | RBF kernel with $\sigma = 0.01$ and $\sigma = 1$            | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |
| 10 | Euclidean                                                    | 69.81    | 78.70       | 85.99     | 81.78    | 0.06         |
|    | 10th polynomial kernel                                       | 83.69    | 94.27       | 88.18     | 91.10    | 2.25         |
|    | RBF kernel with $\sigma = 0.01$ and $\sigma = 1$            | 88.78    | 100.00      | 88.78     | 94.05    | 0.02         |

4. Conclusion
Schizophrenia affects the mind, feeling, and behavior and one of the severe mental disorders worldwide. Earlier treatment for this disease is essential; therefore, fuzzy kernel c-means was proposed in this research. This method was compared with fuzzy kernel c-means and utilize the schizophrenia dataset that was obtained from Northwestern University. There are several kernel functions, but RBF and polynomial kernel function were used in this research with only several kernel parameters evaluated. As the evaluation, k-fold cross-validation with $k = 3, 5, 7,$ and $10$ was used, and each of the performances was analyzed. From the experiments, it was concluded that fuzzy kernel c-means using RBF kernel with $\sigma = 0.01$ and $\sigma = 1$ provides better performance than polynomial kernel and fuzzy c-means with a similar running time as fuzzy c-means.

Furthermore, the behavior of the kernel parameter in the kernel function used in fuzzy c-means can be analyzed and evaluated. The other kernel function can also be considered to examined using the same or different evaluation in fuzzy kernel c-means. Other clustering methods were also accepted to develop in order to improve the performance in predicting the schizophrenia dataset.

Acknowledgment
This research was supported financially by the Ministry of Research, Technology, and Higher Education Republic of Indonesia (KEMENRISTEKDIKTI) with PTUPT 2020 research grant scheme.

References
[1] World Health Organization. https://www.who.int/ (Accessed: 7 September 2019)
[2] Wu Y, Kang R, Yan Y, Gao K, Li Z, Jiang J, Chi X, and Xia L 2018 J. Int. Med. Res. 46(10)
4039-4049

[3] Cho G, Yim J, Choi Y, Ko J, and Lee S 2019 Psychiatry Investig. 16(4) 262-269
[4] Birnbaum M, Rizvi A, Choudhury MD, Ernala S, Cecchi G, and Kane J 2018 Schizophrenia Bulletin 44 S100-S101
[5] Nimkar AV and Kubal DR 2018 4th Int. Conf. on Computer and Information Sciences (Malaysia: IEEE) pp 1-6
[6] Rampisela TV and Rustam Z 2018 IOP Conf. Series: Journal of Physics: Conf. Series. 1108 012044
[7] Rustam Z and Rampisela TV 2018 Int. J. Eng. Technol. 7(4) 6873-6877.
[8] Wang L, Kogan A, Cobia D, Alpert K, Kolansy A, Miller MI, and Marcus D 2013 Front. Neuroinf. 7(25)
[9] Dunn JC 1973 J. Cybern. 3 32–57
[10] Bezdek JC 1981 Pattern recognition with fuzzy objective function algorithms (New York: Plenum Press)
[11] Rustam Z and Talita AS 2015 J. Theor. Appl. Inf. Technol. 81 161-165
[12] Stehman and Stephen V 1997 Remote Sens. Environ. 62 (1) 77-89