HASIL CEK_FSVM with Distance-Based Similarity Measure

by Fsvm With Distance-based Similarity Measure Cek_fsvm With Distance-based Similarity Measure

Submission date: 02-Jul-2021 10:27AM (UTC+0700)
Submission ID: 1614792990
File name: FSVM_with_Distance-Based_Similarity_Measure.docx (99.44K)
Word count: 4065
Character count: 21959

Optimization of Fuzzy Support Vector Machine (FSVM) Performance by Gaussian Membership Function

Tia Nursofiyani, Sugiyarto Surono*
Departement of Mathematics, Ahmad Dahlan University, Yogyakarta, Indonesia

Abstract

Optimization is applied to determine the maximum or minimum value of function. In this research, optimization is applied to fuzzy membership function from Fuzzy Support Vector Machine (FSVM) algorithm. SVM is one of the methods that effectively works in data classification. One of the main problems in this method is its inability to work effectively in high numbers and complex data because of its sensitiveness to outliers and noises. To overcome this problem, this method needs to be combined with fuzzy logic. A suitable selection on the membership function is necessary to provide the effectiveness of the performance of the FSVM algorithm. This research studied the effectiveness of the new method of the membership function optimization based on the distance-based similarity measure with the application of Euclidean distance, Manhattan distance, Chebyshev distance, and Minkowski distance. The objective of this research is to determine the best distance function that able to optimize the FSVM classification process. Four proposed FSVM models and SVM normal are chosen as the comparison references. The result of the study shows that the applied method is effectively able to diminish the noise effect and increase the classification accuracy. FSVM with the Chebhysev distance matrix provides the highest accuracy percentage in 94% and the penalty parameter in 1000.

Keyword: FSVM, Membersship Function Fuzzy, Distance-based Similarity Measure Classification

1. Introduction

Classification is a grouping method based on the characteristics of the classification object. The most advanced and discussed classification method in the last decade is Support Vector Machine (SVM). It occurs because this method has a wide variety of applications and able to provide high performance in the generalization process. In the study conducted by [1], the SVM method, ANN, KNN, fuzzy logic, and RF (Random Forest) are applied for the classification driving model. The result of this research shows that SVM had the highest accuracy level in 96%. On the other study worked by [2] and [3], SVM had also proven to be able to provide high generalization performance and improve the level accuracy better than the other methods. In the real world, the SVM method can be applied in many sectors, such as text categorization [4] [5], speech recognition [6], bioinformatics [7], network security [7], etc.

SVM was introduced by Vapnik in 1995 [8] [9]. The theory of this method was based on structural risk minimization. SVM method is considered as one of the advanced supervised algorithm method that work to determine the classification model or the optimum group of separation from the classification data that trained in certain algorithm to divide the dataset in two or more clusters. Steve Gunn stated that the objective of the SVM method is to find the optimum global solution by mapping the training data to the high dimension space for then the cluster that able to maximize the margin between two classes is determined [10]. Margin is defined as the distance of the support vector towards the hyper-plane. Support vector is the pattern of each of the clusters that have the nearest distance to the hyper-plane.

In complex problems where the data have a high number of parameters, the application of SVM will be hampered because of the occurrence of the noises and outliers, which results in the decrease of the SVM generalization performance [11]. There are many ways to solve this type of problem, one of these is by combining the SVM method with the fuzzy logic [12]-[14]. The previous research applied fuzzy logic with two methods, which are by applying the fuzzy

*Corresponding author. Sugiyarto Surono, Departement of Mathematics, Ahmad Dahlan University, Yogyakarta, Indonesia. E-mail: sugiyarto@math.uad.ac.id.

rules [15] [16] or membership function fuzzy [17]-[22]. These two methods are applied on each of the samples so that the new input sample is able to provide different contributions to diminish the effect of the noises and outliers and improve the generalization performance of SVM classification.

Fuzzy logic is a mathematical method to define obscurity. This fuzzy logic was introduced by Lotfi A. Zadeh in 1965 [23]. The combination of the SVM method and fuzzy logic is called Fuzzy Support Vector Machine (FSVM). One of the most crucial variables applied in FSVM is the membership function used for the clustering process. There are a few approach that can be applied to build the membership function in fuzzy, such as function, intuition, inference, ordering rank, inductive reasoning, neural networks, and genetic algorithm approach.

Euclidean distance is considered as the main criteria to determine the data similarity, which then will be applied to build a membership function. Xiao-Kang Ding (2018) proposed the method of FSVM based on the Euclidean distance. His research applied three methods for build the membership function, which are FSVM-1 that used conventional method as the comparison references and FSVM-2 and FSVM-3 that used the comparison of the distance of the positive and negative samples to the cluster centroid. The difference between FSVM-2 and FSVM-3 occurs in the initial process before the measurement where the sample coordinate in FSVM-3 will be mapped first toward the high-dimension plane. This research concludes that FSVM-3 method gives the best level of accuracy, followed by FSVM-2 and FSVM-1.

The measurement distance is the most common method to determine the similarity measurement. Few distance parameters are usually applied, such as Euclidean distance, Manhattan distance, Hamming distance, Minkowski-Chebyshev distance, etc. Each of these distance parameters applications has its own advantages and disadvantages.

Based on Mohammad and Abdulazeez (2018), Euclidean distance is a distance that is commonly applied to measure the distance on the numerical data. Euclidean works efficiently by measuring the similarity in the cluster and able to separate the data correctly [25]. Manhattan distance is usually applied because of its ability to predict specific parameters such as outlier appearance [26]. Furthermore, Chebyshev distance also has the advantage through its sensitivity towards the object with an outlier.

Built upon the statement in the previously mentioned research, we can conclude that the combination of the SVM method and fuzzy logic is able to optimize the classification process through the suitable selection of the membership function. This research will apply the FSVM method with the Gaussian membership function based on the measures distance. To improve the performance of the FSVM clustering, a comparative study will be applied through a few different measures distance methods such as Euclidean distance, Manhattan distance, Chebyshev distance, Minkowski distance, Hamming distance, and Minkowski-Chebyshev distance.

2. Fuzzy Support Vector Machine (FSVM)

The fuzzy system proposed by Lotfi A. Zadeh is formed based on the theory of fuzzy set and fuzzy logic. This method works for solving complex real-life problems such as uncertainty and inaccuracy. The fuzzy set theory and the clustering technique are considered very important and beneficial in solving the uncertainty and inaccuracy that leads to the improvement of the generalization classification performance. FSVM is an extension method introduced by Lin and Wang in 2002. The function of this method is to decrease the SVM sensitiveness toward outliers and noises. FSVM works by determining the lowest weight for each sample measured by the fuzzy membership function. Furthermore, this fuzzy membership function is measured based on the distance similarity of the data. The fuzzy membership function (s_i) is applied with $0 < s_i \leq 1$ on the training dataset x_i in the class of $y_i \in [1, -1]$. Therefore, the formed fuzzy dataset can be defined as $\{(x_1, y_1, s_1), (x_2, y_2, s_2), \dots, (x_n, y_n, s_n)\}$.

Determine the optimum hyper-plane FSVM by inputting the value of the membership function into the SVM standard formula.

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=0}^{n} s_i \, \xi_i \qquad (1)$$

$$y_i(w^T x_i + b) \geq 1 - \xi_i \,;\, \xi_i \geq 0; \quad 1 \leq i \leq n \qquad (2)$$

Where s_i is the fuzzy membership function ranged from 0 to 1 $(0 < s_i \leq 1)$.

The membership degree for each of the data is measured by adopting the membership function measurement proposed by Xiao-Kang and co. in 2016. This function works by comparing the distance for each of the positive and negative samples towards each of its cluster centroid where the measurement of the

Euclidean distance, Manhattan distance, Chebyshev distance, and Minkowski distance will be applied.

Hence, the calculation of the membership function can be formulated as:

$$s_i = \begin{cases} f(d_i^+), if \|x_i^+ - x_{center}^+\| \geq \|x_i^+ - x_{center}^-\| \\ 1, ifj\ \|x_i^+ - x_{center}^+\| < \|x_i^+ - x_{center}^-\| \\ 1, if \|x_i^- - x_{center}^+\| > \|x_i^- - x_{center}^-\| \\ f(d_i^-), if \|x_i^- - x_{center}^+\| \leq \|x_i^- - x_{center}^-\| \end{cases} \quad (3)$$

Where

$$d_i \begin{cases} \|x_i^+ - x_{center}^+\| \\ \|x_i^- - x_{center}^+\| \end{cases} \quad (4)$$

When the distance of the positive sample to the positive cluster centroid is lower than the distance of this sample to the negative cluster centroid, the sample can be considered as “worked coordinate”, and its membership is set to 1. On the other hand, when the distance of the positive sample to the positive cluster centroid is bigger than to the negative cluster centroid, this coordinate will be considered as “noisy point” and its membership is measured through the Gaussian membership function.

2.1 Gaussian Membership Function

The formula of the Gauss curve membership function can be defined as:

$$G(x_i;\sigma,c) = e^{\frac{-(x_i-c)}{2\sigma^2}} \quad (5)$$

2.2 Kernel Radial Basis Function (RBF)

The formula of Kernel Radial Basis Function can be defined as:

$$K(x,y) = exp\left(\frac{\|x-y\|^2}{2\sigma^2}\right)$$

2.3 Dintance -Based Similarity Measure

In the application of pattern matching or clustering process, Similarity is one of the main component that require to be considered. A Distance-based similarity measure is one of the type of Similarity Measure, which applied for measuring the similarity level of two variables from the term of geometric distance from the variables that included in these two objects. The distance parameters that considered as the distance-based similarity measure are defined as follows:

1. **Euclidean Distance**

Euclidean distance is one of the distance measurements that are commonly applied to determine the data similarity. To measure the similarity level of the data, the following formula needs to be applied:

$$d_{euclidean} = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}, \quad (6)$$

$$i = 1, 2, 3, \ldots, n$$

2. **Manhattan Distance**

Manhattan distance is applied to measure the absolute difference between the coordinate of two objects. The formula of this distance is defined as follows:

$$d_{Manhattan} = \sum_{i=1}^{n}\|x_i - y_i\|, \quad (7)$$

$$i = 1, 2, 3, \ldots, n$$

3. **Chebyshev Distance**

The measurement of this distance is described as follows:

$$d_{Chebyshev} = (max|x_i - y_i|), \quad (8)$$

$$i = 1, 2, 3, \ldots, n$$

4. **Minkowsky Distance**

Minkowski distance is the generalization of the Euclidean and Manhattan distance. The square (p) is defined as the decisive parameter. If p=1, then the distance of the Minkowski distance is equal to the Manhattan distance. In addition to that, if p=2, then the distance will be considered as the Euclidean distance. The formula of the Minkowski distance can be defined as follows:

$$d_{minkowski} = \left(\sum_{i=1}^{n}|x_i - y_i|^p\right)^{\frac{1}{p}}, \quad (9)$$

$$i = 1,2,3, \dots, n$$

3. Result and Discussion

The algorithm of the Fuzzy Support Vector Machine is simulated towards the breast cancer data acquired from https://archive.ics.uci.edu/ml/datasets/Haberman%27s+Survival. The number of the data is 306 data with three variables. The applied variables are the Age of the patients during surgery (X_1), the year of the patient's surgery (X_2), and the detected number of the positive axilla nodes (X_3) with two classes for each variable, which are class 1, the class where the patient survives in 5 or more years, and class -1, the class where the patient dead in 5 years. The data is processed on the Notebook Jupyter software with the Phyton programming language.

The application of the Fuzzy Support Vector Machine is initiated with determining the fuzzy membership degree of the data, to acquire that value, the cluster centroid of each cluster needs to be previously determined. Considering that the applied data for the FSVM measurement has two classes, which are positive class and negative class, the cluster for each class is defined as the average vector of the attributes. The equation for determining the cluster centroid is described as follows:

$$x^{+}_{center} = \frac{1}{n_+} \sum_{i=1}^{n_+} x_i \qquad (9)$$

$$x^{-}_{center} = \frac{1}{n_-} \sum_{i=1}^{n_-} x_i \qquad (10)$$

Where x^{+}_{center} and x^{-}_{center} are the average value of the positive and negative cluster. n_+ is the number of the data coordinates in the positive cluster and n_- is the number of the data coordinates in the negative cluster.

The calculation of the distance matrix from the data coordinates in each cluster towards each of its cluster centroid is worked through the equation (6), (7), (8), and (9). The result of this calculation was then applied to determine the fuzzy membership degree. Based on equation (3), the acquired result is described as follows:

Table 1
Membership function based on Euclidean distance

X_1	X_2	X_3
0.150079	0.514862	0.716328
0.150079	0.633646	0.985813
1.000000	1.000000	1.000000
⋮	⋮	⋮
1.000000	1.000000	1.000000
1.000000	1.000000	1.000000
1.000000	1.000000	1.000000

Table 2
Membership function based on Manhattan distance

X_1	X_2	X_3
0.150079	0.514862	0.716328
1.000000	1.000000	1.000000
1.000000	1.000000	1.000000
⋮	⋮	⋮
0.112199	0.964693	0.954321
0.092367	0.436997	0.892672
0.075413	0.801741	0.999364

Table 3
Membership function based on Chebysev distance

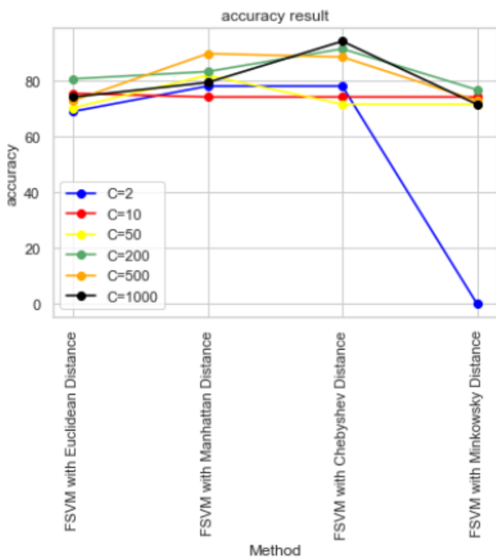
X_1	X_2	X_3
0.150079	0.514862	0.716328
0.150079	0.633646	0.985813
1.000000	1.000000	1.000000
⋮	⋮	⋮
1.000000	1.000000	1.000000
1.000000	1.000000	1.000000
1.000000	1.000000	1.000000

Table 4
Membership function based on Minkowsky distance

X_1	X_2	X_3
0.150079	0.514862	0.716328
0.150079	0.633646	0.985813
1.000000	1.000000	1.000000
⋮	⋮	⋮
1.000000	1.000000	1.000000
1.000000	1.000000	1.000000
1.000000	1.000000	1.000000

The penalty parameter (C) is then applied to the acquired result. There are 7 penalty parameters used in this research, such as $C = 2, C = 10, C = 50, C = 200, C = 500$, $C = 1000$ and kernel RBF. The result of the FSVM clustering process is illustrated in figure 1 below.

Figure 1
The FSVM Accuracy Results



Based on figure 1, we can observe that the FSVM with Manhattan distance and FSVM with Chebyshev distance provides the best clustering performance. The comparison analysis of the FSVM with the pre-determined distance function and SVM can be observed in Figure 2 below.

Figure 2
The Comparison of FSVM and SVM Results

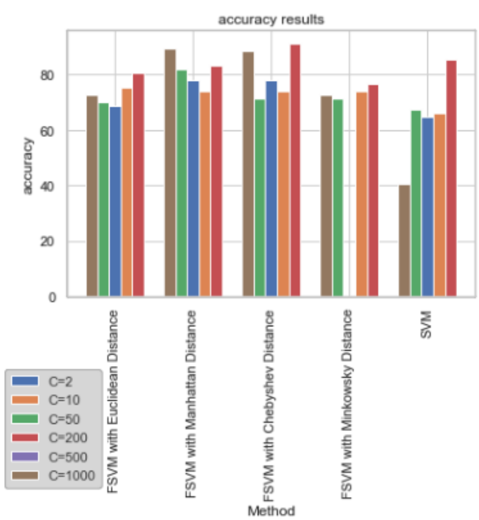


Figure 2 shows that when the penalty parameter is valued at 2 (C=2), the highest accuracy results are produced by FSVM with Manhattan distance and FSVM with Chebyshev distance. Furthermore, in the parameter C = 50 and C = 100, the best accuracy result is acquired by the FSVM with Manhattan distance. As in penalty parameter valued 10, C = 10, FSVM with Euclidean distance provides the best accuracy result. In addition to that, the highest accuracy results from the penalty parameter valued 200 and 1000 is acquired through the FSVM with Chebyshev distance. These accuracy results are illustrated in Table 5 below.

Table 5
The Clustering Accuracy Results

	FSVM with Euclidean	FSVM with Manhattan	FSVM with Chebyshev	FSVM with Minkowsky	SVM
$C = 2$	68.83	77.92	77.92	0.0	64.93
$C = 10$	75.32	74.02	74.03	74.03	66.23
$C = 50$	70.13	81.81	71.43	71.43	67.53
$C = 200$	80.52	83.12	91.30	76.62	85.36
$C = 500$	72.72	89.58	88.37	72.72	40.32
$C = 1000$	74.03	79.22	94	71.42	80.52

From the table 5 above, we can observe that from all of the methods proposed in this research, the FSVM with Chebyshev distance with Kernel RBF and penalty parameter valued 1000 is providing the best performance.

4. Conclusion

In this research, we proposed a combination of the membership function and the Distance-based Similarity Measure with Euclidean, Manhattan, Chebyshev, and Minkowski distance as an effort to determine which distance parameter has the best ability to optimize the Fuzzy Support Vector Machine clustering process. This method can be considered as a newly developed method because there are no previous researchers that analyze the application of FSVM and Distance-based Similarity Result combination.

Based on the result and the discussion of this paper, it can be concluded that the FSVM with Chebyshev distance provides the highest level of performance for the data that considered as imbalanced and full of noise.

In the future, the research about the application of FSVM with Distance-based Similarity Measure for the multivariate data continued with the dimension reduction application needs to be worked for observing how far this type of model can perform effectively and efficiently.

REFERENCES

[1] I. Silva and J. E. Naranjo, "A Systematic Methodology to Evaluate Prediction Models for Driving Style Classification," *Sensors (Switzerland)*, vol. 20, no. 6, pp. 1–21, 2020.

[2] H. Asri, H. Mousannif, H. Al Moatassime, and T. Noel, "Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis," *Procedia Comput. Sci.*, vol. 83, no. Fams, pp. 1064–1069, 2016.

[3] F. Fauzi, "K-Nearset Neighbor (K-NN) dan Support Vector Machine (SVM) untuk Klasifikasi Indeks Pembangunan Manusia Provinsi Jawa Tengah," *J. MIPA*, vol. 40, no. 2, pp. 118–124, 2017.

[4] S. Chatterjee, P. George Jose, and D. Datta, "Text Classification Using SVM Enhanced by Multithreading and CUDA," *Int. J. Mod. Educ. Comput. Sci.*, vol. 11, no. 1, pp. 11–23, 2019.

[5] N. I. Widiastuti, E. Rainarli, and K. E. Dewi, "Peringkasan dan Support Vector Machine pada Klasifikasi Dokumen," *J. Infotel*, vol. 9, no. 4, p. 416, 2017.

[6] S. Pavithra and S. Janakiraman, "Enhanced polynomial kernel (EPK) based support vector machine (SVM) (EPK SVM) classification technique for speech recognition in hearing impaired listeners," *Concurr. Comput.*, no. January, pp. 1–12, 2020.

[7] R. Akbani and T. Korkmaz, "Applications of Support Vector Machines in Bioinformatics and Network Security," *Appl. Mach. Learn.*, no. February 2010, 2010.

[8] V. Vanpik and C. Cortes, "Support Vector Network," *Mach. Learn.*, vol. 20, pp. 273–297, 1995.

[9] P. A. Octaviani, Yuciana Wilandari, and D. Ispriyanti, "Penerapan Metode Klasifikasi Support Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang," *J. Gaussian*, vol. 3, no. 8, pp. 811–820, 2014.

[10] R. G. Brereton and G. R. Lloyd, *Support Vector Machines for classification and regression*. 1998.

[11] Q. Wu and W. Wang, "Piecewise-Smooth Support Vector Machine for Classification," *Math. Probl. Eng.*, vol. 2013, 2013.

[12] C. F. Lin and S. De Wang, "Fuzzy Support Vector Machines," *IEEE Trans. Neural Networks*, vol. 13, no. 2, pp. 464–471, 2002.

[13] W. An and M. Liang, "Fuzzy Support Vector Machine Based on within-Class Scatter for Classification Problems with Outliers or Noises," *Neurocomputing*, 2013.

[14] T. Guernine and K. Zeroual, "New Fuzzy Multi-class Method to Train SVM Classifier," *3rd Int. Conf. Adv. Databases, Knowledge, Data Appl.*, no. c, pp. 77–82, 2011.

[15] J. L. Castro, L. D. Flores-Hidalgo, C. J. Mantas, and J. M. Puche, "Extraction of fuzzy rules from support vector machines," *Fuzzy Sets Syst.*, vol. 158, no. 18, pp. 2057–2077, 2007.

[16] K. Wu, M. Zhou, X. Sean Lu, and L. Huang, "Fuzzy logic based text classification method for social media data," *2017 IEEE Int. Conf. Syst. Man, Cybern. SMC 2017*, vol. 2017-Janua, pp. 1942–1947, 2017.

[17] P. Sridevi, "Identification of Suitable Membership and Kernel Function for FCM Based FSVM Classifier Model," *Cluster Comput.*, vol. 22, no. s5, pp. 11965–11974, 2019.

[18] D. Gupta, B. Richhariya, and P. Borah, "A Fuzzy Twin Support Vector Machine Based on Information Entropy for Class Imbalance Learning," *Neural Comput. Appl.*, vol. 31, no. 11, pp. 7153–7164, 2019.

[19] D. Xiaokang, Y. Lei, Y. Jianping, and Z. Zhaozhong, "Optimization and Analysis on Fuzzy SVM for Objects Classification," *Open Cybern. Syst. J.*, vol. 9, pp. 1421–1429, 2018.

[20] R. Wang, X. Zhang, and W. Cao, "Clifford Fuzzy Support Vector Machines for Classification," *Adv. Appl. Clifford Algebr.*, vol. 26, no. 2, pp. 825–846, 2016.

[21] D. Xiaokang, Y. Lei, Y. Jianping, and Z. Zhaozhong, "Optimization and Analysis on Fuzzy SVM for targets Classification in Forest," *Open Cybern. Syst. J.*, vol. 10, pp. 155–162, 2016.

[22] X. Jiang, Z. Yi, and J. C. Lv, "Fuzzy SVM With A New Fuzzy Membership Function," *Neural Comput. Appl.*, vol. 15, no. 3–4, pp. 268–276, 2006.

[23] N. S. Al-Mumtazah and S. Surono, "Quadratic Form Optimization with Fuzzy Number Parameters: Multiobjective Approaches," *Int. J. Fuzzy Syst.*, vol. 22, no. 4, pp. 1191–1197, 2020.

[24] W. Liu, L. L. Ci, and L. P. Liu, "A New Method of Fuzzy Support Vector Machine Algorithm for Intrusion Detection," *Appl. Sci.*, vol. 10, no. 3, 2020.

[25] N. N. Mohammed and A. M. Abdulazeez, "Evaluation of partitioning around medoids algorithm with various distances on microarray data," *Proc. - 2017 IEEE Int. Conf. Internet Things, IEEE Green Comput. Commun. IEEE Cyber, Phys. Soc. Comput. IEEE Smart Data, iThings-GreenCom-CPSCom-SmartData 2017*, vol. 2018-Janua, pp. 1011–1016, 2018.

[26] U. Syaripudin, I. Badruzaman, E. Yani, D. K, and M. Ramdhani, "Studi Komparatif Penerapan Metode Hierarchical, K-Means Dan Self Organizing Maps (SOM) Clustering Pada Basis Data," *Istek*, vol. VII, no. 1, pp. 132–149, 2013.

[1] I. Silva and J. E. Naranjo, "A Systematic Methodology to Evaluate Prediction Models for Driving Style Classification," *Sensors (Switzerland)*, vol. 20, no. 6, pp. 1–21, 2020.

[2] H. Asri, H. Mousannif, H. Al Moatassime, and T. Noel, "Using Machine Learning Algorithms for Breast Cancer Risk Prediction and Diagnosis," *Procedia Comput. Sci.*, vol. 83, no. Fams, pp. 1064–1069, 2016.

[3] F. Fauzi, "K-Nearset Neighbor (K-NN) dan Support Vector Machine (SVM) untuk Klasifikasi Indeks Pembangunan Manusia Provinsi Jawa Tengah," *J. MIPA*, vol. 40, no. 2, pp. 118–124, 2017.

[4] S. Chatterjee, P. George Jose, and D. Datta, "Text Classification Using SVM Enhanced by Multithreading and CUDA," *Int. J. Mod. Educ. Comput. Sci.*, vol. 11, no. 1, pp. 11–23, 2019.

[5] N. I. Widiastuti, E. Rainarli, and K. E. Dewi, "Peringkasan dan Support Vector Machine pada Klasifikasi Dokumen," *J. Infotel*, vol. 9, no. 4, p. 416, 2017.

[6] S. Pavithra and S. Janakiraman, "Enhanced polynomial kernel (EPK) based support vector machine (SVM) (EPK SVM) classification technique for speech recognition in hearing impaired listeners," *Concurr. Comput.*, no. January, pp. 1–12, 2020.

[7] R. Akbani and T. Korkmaz, "Applications of Support Vector Machines in Bioinformatics and Network Security," *Appl. Mach. Learn.*, no. February, 2010.

[8] V. Vanpik and C. Cortes, "Support Vector Network," *Mach. Learn.*, vol. 20, pp. 273–297, 1995.

[9] P. A. Octaviani, Yuciana Wilandari, and D. Ispriyanti, "Penerapan Metode Klasifikasi Support Vector Machine (SVM) pada Data Akreditasi Sekolah Dasar (SD) di Kabupaten Magelang," *J. Gaussian*, vol. 3, no. 8, pp. 811–820, 2014.

[10] R. G. Brereton and G. R. Lloyd, *Support Vector Machines for classification and regression*. 1998.

[11] Q. Wu and W. Wang, "Piecewise-Smooth Support Vector Machine for Classification," *Math. Probl. Eng.*, vol. 2013, 2013.

[12] C. F. Lin and S. De Wang, "Fuzzy Support Vector Machines," *IEEE Trans. Neural Networks*, vol. 13, no. 2, pp. 464–471, 2002.

[13] W. An and M. Liang, "Fuzzy Support Vector Machine Based on within-Class Scatter for Classification Problems with Outliers or Noises," *Neurocomputing*, 2013.

[14] T. Guernine and K. Zeroual, "New Fuzzy Multi-Class Method to Train SVM Classifier," *3rd Int. Conf. Adv. Databases, Knowledge, Data Appl.*, no. c, pp. 77–82, 2011.

[15] J. L. Castro, L. D. Flores-Hidalgo, C. J. Mantas, and J. M. Puche, "Extraction of fuzzy rules from support vector machines," *Fuzzy Sets Syst.*, vol. 158, no. 18, pp. 2057–2077, 2007.

[16] K. Wu, M. Zhou, X. Sean Lu, and L. Huang, "Fuzzy logic based text classification method for social media data," *2017 IEEE Int. Conf. Syst. Man, Cybern. SMC 2017*, vol. 2017-Janua, pp. 1942–1947, 2017.

[17] P. Sridevi, "Identification of Suitable Membership and Kernel Function for FCM Based FSVM Classifier Model," *Cluster Comput.*, vol. 22, no. s5, pp. 11965–11974, 2019.

[18] D. Gupta, B. Richhariya, and P. Borah, "A Fuzzy Twin Support Vector Machine Based on Information Entropy for Class Imbalance Learning," *Neural Comput. Appl.*, vol. 31, no. 11, pp. 7153–7164, 2019.

[19] D. Xiaokang, Y. Lei, Y. Jianping, and Z. Zhaozhong, "Optimization and Analysis on Fuzzy SVM for Objects Classification," *Open Cybern. Syst. J.*, vol. 9, pp. 1421–1429, 2018.

[20] R. Wang, X. Zhang, and W. Cao, "Clifford Fuzzy Support Vector Machines for Classification," *Adv. Appl. Clifford Algebr.*, vol. 26, no. 2, pp. 825–846, 2016.

[21] D. Xiaokang, Y. Lei, Y. Jianping, and Z. Zhaozhong, "Optimization and Analysis on Fuzzy SVM for targets Classification in Forest," *Open Cybern. Syst. J.*, vol. 10, pp. 155–162, 2016.

[22] X. Jiang, Z. Yi, and J. C. Lv, "Fuzzy SVM

With A New Fuzzy Membership Function," *Neural Comput. Appl.*, vol. 15, no. 3–4, pp. 268–276, 2006.

[23] N. S. Al-Mumtazah and S. Surono, "Quadratic Form Optimization with Fuzzy Number Parameters: Multiobjective Approaches," *Int. J. Fuzzy Syst.*, vol. 22, no. 4, pp. 1191–1197, 2020.

[24] W. Liu, L. L. Ci, and L. P. Liu, "A New Method of Fuzzy Support Vector Machine Algorithm for Intrusion Detection," *Appl. Sci.*, 39. 10, no. 3, 2020.

[25] N. N. Mohammed and A. M. Abdulazeez, "Evaluation of partitioning around medoids algorithm with various distances on microarray data," *Proc. - 2017 IEEE Int. Conf. Internet Things, IEEE Green Comput. Commun. IEEE Cyber, Phys. Soc. Comput. IEEE Smart Data, iThings-GreenCom-CPSCom-SmartData 2017*, vol. 2018-Janua, pp. 1011–1016, 2018.

[26] U. Syaripudin, I. Badruzaman, E. Yani, D. K, and M. Ramdhani, "Studi Komparatif Penerapan Metode Hierarchical, K-Means Dan Self Organizing Maps (SOM) Clustering Pada Basis Data," *Istek*, vol. VII, no. 1, pp. 132–149, 2013.

HASIL CEK_FSVM with Distance-Based Similarity Measure

ORIGINALITY REPORT

17%	11%	14%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	journal.uad.ac.id Internet Source	1%
2	benthamopen.com Internet Source	1%
3	Sugiyarto Surono, Rizki Desia Arindra Putri. "Optimization of Fuzzy C-Means Clustering Algorithm with Combination of Minkowski and Chebyshev Distance Using Principal Component Analysis", International Journal of Fuzzy Systems, 2020 Publication	1%
4	irep.iium.edu.my Internet Source	1%
5	www.researchgate.net Internet Source	1%
6	sci2s.ugr.es Internet Source	1%
7	ejournals.umn.ac.id Internet Source	1%

8	www.ijrte.org Internet Source	1 %
9	Fatima Es-Sabery, Abdellatif Hair, Junaid Qadir, Beatriz Sainz-De-Abajo, Begona Garcia-Zapirain, Isabel Torre-Diez. "Sentence-Level Classification Using Parallel Fuzzy Deep Learning Classifier", IEEE Access, 2021 Publication	1 %
10	C. Victoria Priscilla, D. Padma Prabha. "Influence of Optimizing XGBoost to handle Class Imbalance in Credit Card Fraud Detection", 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT), 2020 Publication	1 %
11	downloads.hindawi.com Internet Source	1 %
12	eprints.uns.ac.id Internet Source	1 %
13	journals.ums.ac.id Internet Source	1 %
14	profdoc.um.ac.ir Internet Source	1 %
15	Rui Wang, Kaili Wang, Wenming Cao, Xiangyang Wang. "Geometric Algebra in Signal	<1 %

and Image Processing: A Survey", IEEE Access, 2019

Publication

16	ijssst.info Internet Source	<1 %
17	Youness Madani, Mohammed Erritali, Jamaa Bengourram, Francoise Sailhan. "A multilingual fuzzy approach for classifying Twitter data using fuzzy logic and semantic similarity", Neural Computing and Applications, 2019 Publication	<1 %
18	Liming Fang, Yang Li, Zhe Liu, Changchun Yin, Minghui Li, Zehong Jimmy Cao. "A Practical Model Based on Anomaly Detection for Protecting Medical IoT Control Services Against External Attacks", IEEE Transactions on Industrial Informatics, 2021 Publication	<1 %
19	Shengnan Wang, Xiaoyong Zhang, Yijun Cheng, Fu Jiang, Wentao Yu, Jun Peng. "A Fast Content-Based Spam Filtering Algorithm with Fuzzy-SVM and K-means", 2018 IEEE International Conference on Big Data and Smart Computing (BigComp), 2018 Publication	<1 %
20	Ricardo Alfredo Cajo Diaz, Mihaela Ghita, Dana Copot, Isabela Roxana Birs, Cristina	<1 %

Muresan, Clara Ionescu. "Context Aware Control Systems: An Engineering Applications Perspective", IEEE Access, 2020
Publication

21 www.math.sci.kmitl.ac.th <1 %
Internet Source

22 B. Richhariya, M. Tanveer. "A robust fuzzy least squares twin support vector machine for class imbalance learning", Applied Soft Computing, 2018 <1 %
Publication

23 Submitted to Universiti Teknologi MARA <1 %
Student Paper

24 Mohammad Daoud. "Novel Approach towards Arabic Question Similarity Detection", 2019 2nd International Conference on new Trends in Computing Sciences (ICTCS), 2019 <1 %
Publication

25 Ren, Y.. "A parsimony fuzzy rule-based classifier using axiomatic fuzzy set theory and support vector machines", Information Sciences, 20111201 <1 %
Publication

26 Annisa Eka Haryati, Sugiyarto. "Clustering with Principal Component Analysis and Fuzzy Subtractive Clustering Using Membership Function Exponential and Hamming Distance", <1 %

IOP Conference Series: Materials Science and Engineering, 2021
Publication

---

27 Hong-Sen Yan, Duo Xu. "An Approach to Estimating Product Design Time Based on Fuzzy ν-Support Vector Machine", IEEE Transactions on Neural Networks, 2007
Publication
<1 %

---

28 Olfa Mabrouk, Lobna Hlaoua, Mohamed Nazih Omri. "Fuzzy Twin SVM based-Profile Categorization approach", 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2018
Publication
<1 %

---

29 Qing Wu, Wenqing Wang. "Piecewise-Smooth Support Vector Machine for Classification", Mathematical Problems in Engineering, 2013
Publication
<1 %

---

30 Studies in Fuzziness and Soft Computing, 2015.
Publication
<1 %

---

31 Wenjuan An, Mangui Liang, He Liu. "An improved one-class support vector machine classifier for outlier detection", Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, 2014
<1 %

Publication

32	Ying Chen. "Efficient Text Classification by Weighted Proximal SVM", Fifth IEEE International Conference on Data Mining (ICDM 05), 2005 Publication	<1 %
33	link.springer.com Internet Source	<1 %
34	moam.info Internet Source	<1 %
35	scholarworks.bridgeport.edu Internet Source	<1 %
36	www.mecs-press.org Internet Source	<1 %
37	Ai-bing Ji, Songcan Chen, Qiang Hua. "Fuzzy classifier based on fuzzy support vector machine", Journal of Intelligent & Fuzzy Systems, 2014 Publication	<1 %
38	Nouh Sabri Elmitwally. "Building A Multi-class XGBoost Model for Arabic Figurative Language", 2020 2nd International Conference on Computer and Information Sciences (ICCIS), 2020 Publication	<1 %
39	hdl.handle.net Internet Source	

		<1%
40	Li, Kai, and Xiaoxia Lu. "A Double Margin Based Fuzzy Support Vector Machine Algorithm", Journal of Computers, 2011. Publication	<1%
41	Humaid Alshamsi, Veton Kepuska, Hazza Alshamsi, Hongying Meng. "Automated Facial Expression and Speech Emotion Recognition App Development on Smart Phones using Cloud Computing", 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2018 Publication	<1%
42	eprints.uad.ac.id Internet Source	<1%
43	dblp.uni-trier.de Internet Source	<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		