Biased support vector machine and weighted-smote in handling class imbalance problem

Hartono
STMIK IBBI

Opim Salim Sitompul, Tulus, Erna Budhiarti Nababan
Universitas Sumatera Utara

International Journal of Advances in Intelligent Informatics Vol. 4 No. 1, Maret 2018

1. Submit pada OJS Jurnal : 14 Januari 2018
2. Accepted with Major Revision: 6 Maret 2018
3. Submit hasil revisi : 11 Maret 2018
4. Editor Decision Accepted: 13 Maret 2018

Bukti Korespondensi Pengajuan Guru Besar Hartono

Judul Karya Ilmiah	:	Biased support vector machine and weighted-smote in handling class imbalance problem
Jurnal	:	International Journal of Advances in Intelligent Informatics
Penulis	:	Hartono, Opim Salim Sitompul, Tulus, Erna Budhiarti Nababan
Edisi	:	2018
Volume	:	Vol. 4 No. 1
Penerbit	:	Universitas Ahmad Dahlan

Home > Vol 4, No 1 (2018) > **Hartono**

Biased support vector machine and weighted-smote in handling class imbalance problem

(1) * Hartono Hartono *(Universitas Sumatera Utara, Indonesia)*
(2) Opim Salim Sitompul *(Universitas Sumatera Utara, Indonesia)*
(3) Tulus Tulus *(Universitas Sumatera Utara, Indonesia)*
(4) Erna Budhiarti Nababan *(Universitas Sumatera Utara, Indonesia)*
* *corresponding author*

Check for updates

Abstract

Class imbalance occurs when instances in a class are much higher than in other classes. This machine learning major problem can affect the predicted accuracy. Support Vector Machine (SVM) is robust and precise method in handling class imbalance problem but weak in the bias data distribution, Biased Support Vector Machine (BSVM) became popular choice to solve the problem. BSVM provide better control sensitivity yet lack accuracy compared to general SVM. This study proposes the integration of BSVM and SMOTEBoost to handle class imbalance problem. Non Support Vector (NSV) sets from negative samples and Support Vector (SV) sets from positive samples will undergo a Weighted-SMOTE process. The results indicate that implementation of Biased Support Vector Machine and Weighted-SMOTE achieve better accuracy and sensitivity.

Keywords

Class Imbalance; Biased Support Vector Machine; Borderline-SMOTE; Positive Samples; Negative Samples

\

Bukti Submit pada OJS Jurnal

#146 Summary

SUMMARY REVIEW EDITING

Submission

Authors	Hartono Hartono, Opim Salim Sitompul, Tulus Tulus, Erna Budhiarti Nababan
Title	Biased support vector machine and weighted-smote in handling class imbalance problem
Original file	146-530-1-SM.DOCX 2018-01-14
Supp. files	None
Submitter	Mr. Hartono -
Date submitted	January 14, 2018 - 09:46 PM
Section	Articles
Editor	Mohammad Syafrullah
Abstract Views	6700

Notifikasi Registrasi User Baru pada Gmail

[SAIN 2018] New User Registration   Kotak Masuk ×

Andri Pranolo <andri.pranolo.id@ieee.org>
kepada saya

Thank you for registering as a user with SAIN. Please keep track of your username and password, which are needed for all work with this conference.

Username: hartono
Password: ab

Thank you,
Andri Pranolo

International Symposium on Advanced Intelligent Informatics (SAIN) The 2018 International Symposium on Advanced Intelligent Informatics (SAIN)
http://sain.ijain.org/submit/index.php/sain/sain2018/index

Accepted With Major Revision pada Gmail

Gmail

Hartono Ibbi <hartonoibbi@gmail.com>

[IJAIN] Editor Decision

Andri Pranolo <andri.pranolo@tif.uad.ac.id> 6 Maret 2018 pukul 11.00
Kepada: "Mr. Hartono -" <hartonoibbi@gmail.com>

Mr. Hartono -:

We have reached a decision regarding your submission to International Journal of Advances in Intelligent Informatics, "Implementation of Biased Support Vector Machine and Weighted-SMOTE in Handling Class Imbalance Problem".

Our decision is: Accept with Major Revision

Please kindly submit the revision before March 14th, 2018.

Regards

Andri Pranolo
(Managing Editor)

Reviewer B:

Significance:
- How important is the work reported? Does it attack an important/difficult problem (as opposed to a peripheral/simple one)?
- Does the approach offered advance the state of the art?
- Does it involve or synthesize ideas, methods, approaches from multiple disciplines?
- Does it have interesting implications for multiple disciplines?:
Good

Originality: - Is this a new issue? Is this a novel approach to an issue? - Is this a novel combination of familiar ideas/techniques/methods/approaches? - Does the paper point out differences from related research? - Does the paper properly situate itself with respect to previous work?:
Fair

Quality: - Is the paper technically sound? How are its claims backed up? - Does it carefully evaluate the strengths and limitations of its contribution?:
Fair

Clarity: - Is the paper clearly written? Does it motivate the research? Does it describe clearly the methods employed (e.g., experimental procedures, algorithms, analytical tools), if any? - Are the results, if any, described and evaluated thoroughly? - Is the paper organized in a sensible and logical fashion?:
Good

Relevance:
- Is the paper closely related to the theme of the journal (broadly conceived)?
- Is the content interesting enough to a broad audience?
- Is the paper readable in a multi-disciplinary context?:
Good

Technical (1): Structure of the paper:
Good

Technical (2): Standard of English:
Good

Technical (3): Appropriateness of abstract as a description of the paper:
Good

Technical (4): Use and number of keywords/key phrases:
Good

Technical (5): Relevance and clarity of drawings, graphs and tables:
Good

Technical (6): Discussion and conclusions:
Good

Technical (7): Reference list, adequate and correctly cited:
Good

Explanations for the above ratings and other general comments on major issues:

Comments on the minor details of the article:
1. The class imbalance problem occurs when the clustering results show that there is a class with a much larger number of instances than the other classes. ---> It should be classification and NOT clustering, right?

2. There are several bias SVM methods that had been implemented in handling class imbalance problem. One example is :
https://www.sciencedirect.com/science/article/pii/S1568494613004420

In this paper, NO REVIEW had been done on previous publication in similar topic. I suggest that in Chapter II, related works are discussed.

3. Improvement that can be done :
Read related articles about biased SVM for handling class imbalance problem,

and try to find what the drawback of those previous works. Or what problem
do exist in previous methods. Try to address solution for that problem.

Reviewer F:

Significance:
 - How important is the work reported? Does it attack an
important/difficult problem (as opposed to a peripheral/simple one)?
 - Does the approach offered advance the state of the art?
 - Does it involve or synthesize ideas, methods, approaches from multiple
  disciplines?
 - Does it have interesting implications for multiple disciplines?:
        Fair

Originality: - Is this a new issue? Is this a novel approach to an issue? -
Is this a novel combination of familiar ideas/techniques/methods/approaches?
- Does the paper point out differences from related research? - Does the
paper properly situate itself with respect to previous work?:
        Fair

Quality: - Is the paper technically sound? How are its claims backed up? -
Does it carefully evaluate the strengths and limitations of its
contribution?:
        Poor

Clarity: - Is the paper clearly written? Does it motivate the research? Does
it describe clearly the methods employed (e.g., experimental procedures,
algorithms, analytical tools), if any? - Are the results, if any, described
and evaluated thoroughly? - Is the paper organized in a sensible and logical
fashion?:
        Poor

Relevance:
 - Is the paper closely related to the theme of the journal (broadly
conceived)?
 - Is the content interesting enough to a broad audience?
 - Is the paper readable in a multi-disciplinary context?:
        Fair

Technical (1): Structure of the paper:
        Poor

Technical (2): Standard of English:
        Poor

Technical (3): Appropriateness of abstract as a description of the paper:
        Fair

Technical (4): Use and number of keywords/key phrases:
        Good

Technical (5): Relevance and clarity of drawings, graphs and tables:
        Fair

Technical (6): Discussion and conclusions:
        Poor

Technical (7): Reference list, adequate and correctly cited:
        Fair

Explanations for the above ratings and other general comments on major
issues:
        In order to show the advantages of the proposed methods, the authors should
compare the existing or common methods versus the proposed methods. Analyze
the results and show the improvements.

Comments on the minor details of the article:
        The structure of the presentation should be improved to show the
contribution of the paper.
English grammar needs improvement.

PERBAIKAN YANG DILAKUKAN SESUAI KOMENTAR DARI REVIEWER B

Biased Support Vector Machine and Weighted-SMOTE in Handling Class Imbalance Problem

Hartono[a,b,1,*], Opim Salim Sitompul[b,2], Tulus[b,3], Erna Budhiarti Nababan[b,4]

[a] *STMIK IBBI, Jl. Sei Deli No. 18 Medan, Medan 20114, Indonesia*
[b] *Universitas Sumatera Utara, Jl. Dr. T. Mansyur No. 9, Medan 20155, Indonesia*
[1] *hartonoibbi@gmail.com*; [2] opim@usu.ac.id; [3] tulus_jp@yahoo.com; [4] ernastp@yahoo.com*
* *hartonoibbi@gmail.com*

ARTICLE INFO

Article history:
Received
Revised
Accepted

ABSTRACT

Class imbalance is a situation where instances in one class are much higher than instances in other classes. The problem of class imbalance is a major problem in machine learning because it can affect predictive accuracy. SVM is robust and precise in handling class imbalance problem but cannot handling the class imbalance problem with the bias in the data distribution is significant. There are some correction with the determining process of ideal Hyperplane In SVM with the respecting to the bias. In the some case, if the focus would be placed on the majority class, one of the Method is Biased Support Vector Machine (BSVM). BSVM will give the better sensitivity control compare to SVM but has the lack in accuracy and this problem will overcome using the integration with another approach in Data Level-Solutions.BSVM has been combined with SMOTEBoost in handling class imbalance problem. SMOTEBoost should also respect to all the minority data samples and for this reason, Weighted-SMOTE has been introduced. NSV Sets from negative samples and SV Sets from Positive Samples will then undergo a Weighted-SMOTE process. Performance evaluation in imbalanced domains are accuracy and sensitivity. The results of this study indicate that the implementation of Biased Support Vector Machine and Weighted-SMOTE will achieve better accuracy and sensitivity.

Keywords:
Class Imbalance
Biased Support Vector Machine
Weighted-SMOTE
Accuracy
Sensitivity

I. Introduction

The class imbalance problem occurs when the classification results show that there is a class with a much larger number of instances than the other classes. This raises the issue of majority versus minority [1]. Most of the traditional machine learning techniques deal with reasonably balanced class distributions in solving classification problems [2]. The problem of class imbalance is a major problem in machine learning because it can affect predictive accuracy, where machine learning will provide better accuracy of predicted results for classes with larger instances of instances and the resulting accuracy results will be lower in the class with small number of instances [3]. This problem also increase the number of missclassification in machine learning [4]. In General, there are three types of approaches that can be used to solve class imbalance problems. The first approach is Data-Level Solutions, the second approach is Algorithm-Level Solutions, and the third approach is Cost-Sensitive Solutions [5]. The problem of class imbalance can be divided into two scenarios: binary and multi-class [6].

On completion with each approach it has been suggested that the representativeness of the sample problem is important to determine because it involves three things: it must have a reduced size compared to the original source, containing the main information in source, and low redundancy [7]. SVM is robust and precise, but can be sensitive to missing values and difficult to train for large scale data and highly imbalance datasets. For train large scale data, it must be combined with another approach in Data Level-Solutions [8]. A widely used method is Synthetic Minority Over-sampling TEchnique (SMOTE) method [9]. For controlling the sensitivity in Support Vector Machine, [10] has

Comment [Office1]: The class imbalance problem occurs when the clustering results show that there is a class with a much larger number of instances than the other classes. ---> It should be classification and NOT clustering been introduced Biased Support Vector Machine. Biased Support Vector Machine was combined with SMOTE by [11] in handling class imbalance problem.

Comment [Office2]: There are several bias SVM methods that had been implemented in handling class imbalance problem. One example is : https://www.sciencedirect.com/science/article/pii/S1568494613004420

The SMOTE method will generally initialise a sample number of minorities and after that it will generate equal numbers of synthetic data, but this method should also respect to all the remaining minority data samples [12]. A Weighted-SMOTE proposed by [12] will oversampling of each minority data sample is carried out based on the weight assigned to it. These weights are determined by using the Euclidean distance of a particular minority data sample with respect to all the remaining minority data samples.

This paper introduced the implementation of Biased Support Vector Machine and Weighted-SMOTE in Handling Class Imbalance Problem. Biased Support Vector Machine will give the better sensitivity control compare to Support Vector Machine. Biased Support Vector Machine works by providing different cost functions for positive samples (minority class) and negative samples (majority class) will group both positive samples and negative samples into SV Sets and NSV Sets.

Biased Support Vector Machine has been combined with SMOTEBoost in handling class imbalance problem to get the better accuracy. SMOTEBoost should also respect to all the minority data samples and for this reason, Weighted-SMOTE has been introduced. NSV Sets from negative samples and SV Sets from Positive Samples will then undergo a Weighted-SMOTE process. The traditional performance evaluation in imbalanced evaluation are accuracy and sensitivity [3].

The main contribution in this paper is that once the case if the focus would be placed on the majority class, one of the Method is Biased Support Vector Machine (BSVM) and how to get the better sensitivity and accuracy using BSVM with the implementation of BSVM and Weighted-SMOTE in handling class imbalance problem.

II. Biased Support Vector Machine and Weighted-SMOTE Algorithm

Traditional Support Vector Machine (SVM) cannot handling the class imbalance problem with the bias in the data distribution is significant, it is because the separation hyperplane learned by the SVM is very close to the minority class which explains why SVM has a degrading performance on highly imbalanced datasets [13]. A Geometric Mean SVM (GSVM) proposed by [14] will get the ideal hyperplane by calculating the moving of the original bias in the SVM to improve the geometric mean between the true positive rate and the true negative rate. But in the some problem, The focus of interest is given more to the positive class since it may contain unusual behaviour differs from the general access pattern and to deal with datasets where one class, the positive class, is considered more relevant than another class in binary classification problems, the Biased Support Vector Machine (BSVM) method [10] can be considered.

BSVM is designed for the case when it is non-critical to increase the true positive ratio in exchange for an increase in the false positive rate[15]. BSVM Method will achieve better performance in sensitivity, but have a significant reduction in accuracy[15] and this problem will overcome using the integration with Weighted-Smote Algorithm. Biased Support Vector Machine has been combined with SMOTEBoost in handling class imbalance problem[11] but this method should also respect to all the remaining minority data samples [12] and this is the reason why Weighted-SMOTE has and advantage to use in this reasearch.

Comment [Office3]: ead related articles about biased SVM for handling class imbalance problem, and try to find what the drawback of those previous works. Or what problem do exist in previous methods. Try to address solution for that problem.

In Biased Support Vector Machine, The minority samples are given with a larger cost function and the mathematical model of Biased Support Vector Machine is defined by [10]

$$min \frac{1}{2}|w|^2 + C^+ \sum_{i \in I_+} \xi_i^k + C^- \sum_{i \in I_-} \xi_i^k$$
$$s.t. y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i \quad \xi_i \geq 0, \forall i \qquad (1)$$

Where w and b represent the normal vector and the intercept of hyperplane, respectively $\xi_i \geq 0$ are slack variable. The detailed procedure of Biased Support Vector Machine and Weighted-SMOTE Algorithm in handling class imbalance is as follows.

Step 1, Take some data from minority class and majority class. Some data from the minority class are expressed as positive samples and some data from the majority class is expressed as negative samples.

Step 2, Determine the hyperplane combination of positive samples and negative samples. Where the combined value of positive samples is 1 and the combined value of negative samples is -1.

Step 3, Hyperplane is obtained by using Biased Support Vector Machine which is done by minimizing the margin value of positive samples and negative samples using equation (2) and (3) [11]

$$\frac{1}{2}\|W\|^2 = \frac{1}{2}(w_1^2 + w_2^2 + w_3^2 + w_4^2) \quad (2)$$

s.t.

$$y_i(w_i x_i + b) \geq 1, \qquad i = 1,2,3,\cdots,N \quad (3)$$

$$y_i(w_1.x_1 + w_2.x_2 + w_3.x_3 + w_4.x_4 + b) \geq 1$$

Step 4, Equation (4) is an equation for the determination of Positive Samples and Negative Samples [11]

$$f(x) = sign\left(\sum_{i=1}^{l} y_i(w.x_i) + b\right) \quad (4)$$

If the sign function gives a result greater than 0 then it is included in the minority class and if the sign function gives a result smaller than 0 then it is included in the majority class.

Conduct training for minority class and majority class based on hyperplane equation obtained. Training process can be done using (4)

Step 5, The result for minority class must be positive and the result for majority class must be negative. If the training result for an instance in the minority class gives a negative result then move the instance into the majority class and if the training result for an instance in the majority class gives a positive result then move that instance into the minority class.

Step 6, Take some data back from the minority class and make it as positive samples and retrieve some data back from majority class and make it as negative samples. Determine the hyperplane of each positive samples and negative samples using (2) and (3).

Step 7, Perform training process on minority class by using (5). for the process of determining SV Sets and NSV Sets by using hyperplane from positive samples. If the calculation result gives then it is categorized into SV Sets and otherwise categorized as NSV Sets.

$$H_1 = SV\ Sets\ for\ Minority\ Class\ (w.x\ +\ b = 1) \quad (5)$$

Step 8, Do the training process on majority class by using (6). for the process of determining SV Sets and NSV Sets by using hyperplane of negative samples. If the calculation result gives then it is categorized into SV Sets and otherwise categorized as NSV Sets.

$$H_2 = SV\ Sets\ for\ Majority\ Class\ (w.x\ +\ b = -1) \quad (6)$$

Step 9, SV Sets on the Minority Class will be eliminated noise and then will process the Weighted-SMOTE process to then be combined with NSV Sets on the minority class to become a new minority class. The Weighted-SMOTE process is as follows [12]

(1) Calculate euclidean distance for the each of the *T* minority data samples using (7) [16]

$$d_{ij} = \sqrt{\sum_{k=1}^{p}(x_{ik} - x_{jk})^2} \quad (7)$$

Here, $i = [1,2,\ldots.,T]$ and $j = [1,2,\ldots,T]$ and $j \neq 1$. The *Euclidean Distance* for all the minority data are calculated and stored in a column matrix using (8) [10]

$$ED = [ED_1, ED_2, \ldots, ED_T] \quad (8)$$

(2) Normalized the ED matrix using the maximum of the ED (ED_{max}) and the minimum of ED (ED_{min}) and called it as normalized ED (NED) that can be seen in (9).

$$NED_i = \frac{ED_i - ED_{min}}{ED_{max} - ED_{min}} \quad (9)$$

(3) Convert the *NED* matrix using the (10) to get the Remodeled Normalized Euclidean Distance Matrix (*RNED*)

$$[RNED]_{Tx1} = sum(NED) - [NED]_{Tx1} \quad (10)$$

(4) Calculated the Weight Matrix for each minority of *T* samples with respect to the sum of RNED Matrix using (11)

$$[Weight\ Matrix]_{Tx1} = \frac{[RNED]_{Tx1}}{sum(RNED)} \quad (11)$$

(5) Using this Weight matrix to find the SMOTE generation matrix using (12)

$$[SMOTE\ Generation\ Matrix]_{Tx1} = \frac{NxT}{100}[Weight\ Matix]_{Tx1} \quad (12)$$

Step 10, NSV Sets on Majority Class will experience Weight-SMOTE process and will then be combined with SV Sets on Majority Class to become the new majority class. The Weighted-SMOTE process is same with *Step* 9.

III. Method

The experiments are conducted using *Iris Dataset* and *Balanced Scale Weight & Distance Database* and the software is based on R Language. The Research Method can be seen in Fig. 1

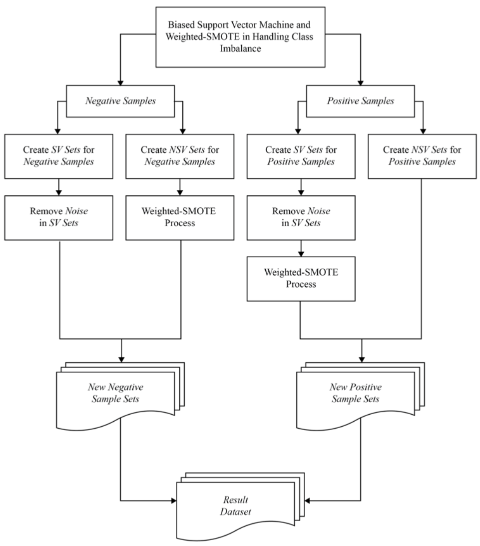

Fig. 1. Research Method

From the Fig. 1, it can be seen that the Biased Support Vector Machine will classify both majority class (negative samples) and minority class (positive samples) into Support Vector Sets (SV Sets) and Non Support Vector Sets (NSV Sets). The noise from the SV Sets Negative Samples will be removed and the NSV Sets Negative Samples will be processed using Weighted-SMOTE. Then SV Sets Negative Samples will combine with NSV Sets Negative Samples, will make a new Negative Samples Sets. After this process, the Noise from the SV Sets in Positive Samples will be removed and this SV Sets will be processed using Weighted-SMOTE. Then SV Sets Positive Samples will combine with NSV Sets Positive Samples, will make a new Positive Samples Sets. New Negative Sample Sets and New Positive Sample Sets will combine to make Result Dataset.

This research will calculate the Accuracy and Sensitivity from the result of implementation of Biased Support Vector Machine and Weighted-SMOTE using (13) [12] and (14) [10].

$$Accuracy = \frac{TN+TP}{TN+FP+FN+TP} \quad (13)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (14)$$

Using Confusion Matrix [10] for determine the value of TN, TP, FN, and FP that can be seen inf Table 1.

Table 1. Confusion Matrix

	Classified as Positive	Classified as Negative
Positive Samples	*True Positive* (TP)	*False Negative* (FN)
Negative Samples	*False Positive* (FP)	*True Negative* (TN)

IV. Results and Discussion

The testing are conducted using *Iris Dataset* and *Balanced Scale Weight & Distance Database* and the software is based on R Language. The result of the testing using *Iris Dataset* can be seen in Table 2.

Table 2. The Result Testing using *Iris Dataset*

Iteration	Before Processing		After Processing		Accuracy	Sensitivity
	Number of Majority	Number of Minority	Number of Majority	Number of Minority		
1	60	40	52	48	0.87	0.83
2	60	34	49	45	0.81	0.79
3	62	39	52	49	0.82	0.79
4	64	36	50	50	0.902	0.904
5	55	44	50	49	0.93	0.92
6	71	29	50	50	0.84	0.83
7	66	30	50	46	0.93	0.92
8	62	37	50	49	0.89	0.88
9	67	35	54	48	0.94	0.93
10	65	35	52	48	0.89	0.88
Average	63.2	35.9	50.9	48.2	0.88	0.87

The result of the testing using *Balanced Scale Weight & Distance Database Dataset* can be seen in Table 3.

Table 3. The Result Testing using *Balanced Scale Weight & Distance Database*

Iteration	Before Processing		After Processing		Accuracy	Sensitivity
	Number of Majority	Number of Minority	Number of Majority	Number of Minority		
1	325	251	290	286	0.83	0.832
2	353	226	291	288	0.84	0.82
3	345	242	301	286	0.91	0.906
4	363	213	295	281	0.87	0.85
5	347	232	291	288	0.83	0.84
6	319	257	292	284	0.88	0.87
7	321	256	294	283	0.9	0.91
8	316	260	293	283	0.86	0.85
9	343	236	290	289	0.79	0.8
10	361	219	294	286	0.85	0.84
Average	339.3	239.2	293.1	285.4	0.86	0.85

From the result that can be seen in Table 2 dan Table 3, there are an improvement in the different of number of majority and number of minority before and after processing with Biased Support Vector Machine and Weighted-SMOTE. The accuracy and sensitivity that has been achieved from the testing is very satisfied. The high value of Accuracy and Sensitivity means that the implementation of Biased Support Vector Machine and Weighted-SMOTE can handle the class imbalance problem.

V. Conclusion

This paper has introduced the implementation of Biased Support Vector Machine and Weighted-SMOTE in handling class imbalance problem. The implementation is conducted using Iris Dataset and Balanced Scale Weight & Distance Database. The result of this testing has given the very satisfied result that indicate in value of the Accuracy and Sensitivity. for future research should increase the number of datasets involved in the test so that the results obtained can be more accurate.

References

[1] S. M. A. Elrahman and A. Abraham, “A Review of Class Imbalance Problem,” *J. Netw. Innov. Comput.*, vol. 1, pp. 332–340, 2013.
[2] J. Ortigosa-Hernández, I. Inza, and J. A. Lozano, “Measuring the class-imbalance extent of multi-class problems,” *Pattern Recognit. Lett.*, vol. 98, pp. 32–38, 2017.
[3] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, “A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches,” *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 42, no. 4, pp. 463–484, Jul. 2012.
[4] Hartono, O. S. Sitompul, Tulus, and E. B. Nababan, “Optimization Model of K-Means Clustering Using Artificial Neural Networks to Handle Class Imbalance Problem,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 288, p. 012075, Jan. 2018.
[5] W.-C. Lin, C.-F. Tsai, Y.-H. Hu, and J.-S. Jhang, “Clustering-Based Undersampling in Class-Imbalanced Data,” *Inf. Sci.*, vol. 409–410, pp. 17–26, May 2017.
[6] J. A. Sáez, B. Krawczyk, and M. Woźniak, “Analyzing the oversampling of different classes and types of examples in multi-class imbalanced datasets,” *Pattern Recognit.*, vol. 57, pp. 164–178, Sep. 2016.
[7] D. Furundzic, S. Stankovic, S. Jovicic, S. Punisic, and M. Subotic, “Distance based resampling of imbalanced classes: With an application example of speech quality assessment,” *Eng. Appl. Artif. Intell.*, vol. 64, pp. 440–461, Sep. 2017.

[8] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, and G. Bing, "Learning from Class-Imbalanced Data: Review of Methods and Applications," *Expert Syst. Appl.*, vol. 73, pp. 220–239, May 2017.
[9] L. Zhang, C. Zhang, R. Gao, R. Yang, and Q. Song, "Using the SMOTE technique and hybrid features to predict the types of ion channel-targeted conotoxins," *J. Theor. Biol.*, vol. 403, pp. 75–84, 2016.
[10] K. Veropoulos, C. Campbell, and N. Cristianini, "Controlling the Sensitivity of Support Vector Machines," presented at the International Joint Conference on AI, 1999, pp. 55–60.
[11] C. Jian, J. Gao, and Y. Ao, "A new sampling method for classifying imbalanced data based on support vector machine ensemble," *Neurocomputing*, vol. 193, pp. 115–122, Jun. 2016.
[12] M. R. Prusty, T. Jayanthi, and K. Velusamy, "Weighted-SMOTE: A Modification to SMOTE for Event Classification in Sodium Cooled Fast Reactors," *Prog. Nucl. Energy*, vol. 100, pp. 355–364, 2017.
[13] H. He and A. Ghodsi, "Rare class classification by support vector machine," in *Pattern Recognition (ICPR), 2010 20th International Conference on*, 2010, pp. 548–551.
[14] L. Gonzalez-Abril, H. Nuñez, C. Angulo, and F. Velasco, "GSVM: An SVM for handling imbalanced accuracy between classes inbi-classification problems," *Appl. Soft Comput.*, vol. 17, pp. 23–31, Apr. 2014.
[15] L. Gonzalez-Abril, C. Angulo, H. Nuñez, and Y. Leal, "Handling binary classification problems with a priority class by using Support Vector Machines," *Appl. Soft Comput.*, vol. 61, pp. 661–669, Dec. 2017.
[16] R. Loohach and K. Garg, "Effect of Distance Functions on K-Means Clustering Algorithm," *Int. J. Comput. Appl.*, vol. 49, no. 6, pp. 7–9, 2012.

PERBAIKAN YANG DILAKUKAN SESUAI KOMENTAR DARI REVIEWER F

Biased Support Vector Machine and Weighted-SMOTE in Handling Class Imbalance Problem

Hartono[a,b,1,*], Opim Salim Sitompul[b,2], Tulus[b,3], Erna Budhiarti Nababan[b,4]

[a] *STMIK IBBI, Jl. Sei Deli No. 18 Medan, Medan 20114, Indonesia*
[b] *Universitas Sumatera Utara, Jl. Dr. T. Mansyur No. 9, Medan 20155, Indonesia*
[1] *hartonoibbi@gmail.com*; [2] opim@usu.ac.id; [3] tulus_jp@yahoo.com; [4] ernastp@yahoo.com*
** hartonoibbi@gmail.com*

ARTICLE INFO

Article history:
Received
Revised
Accepted

Keywords:
Class Imbalance
Biased Support Vector Machine
Weighted-SMOTE
Accuracy
Sensitivity

ABSTRACT

Class imbalance is a situation where instances in one class are much higher than instances in other classes. The problem of class imbalance is a major problem in machine learning because it can affect predictive accuracy. SVM is robust and precise in handling class imbalance problem but cannot handling the class imbalance problem with the bias in the data distribution is significant. There are some correction with the determining process of ideal Hyperplane In SVM with the respecting to the bias. In the some case, if the focus would be placed on the majority class, one of the Method is Biased Support Vector Machine (BSVM). BSVM will give the better sensitivity control compare to SVM but has the lack in accuracy and this problem will overcome using the integration with another approach in Data Level-Solutions.BSVM has been combined with SMOTEBoost in handling class imbalance problem. SMOTEBoost should also respect to all the minority data samples and for this reason, Weighted-SMOTE has been introduced. NSV Sets from negative samples and SV Sets from Positive Samples will then undergo a Weighted-SMOTE process. Performance evaluation in imbalanced domains are accuracy and sensitivity. The results of this study indicate that the implementation of Biased Support Vector Machine and Weighted-SMOTE will achieve better accuracy and sensitivity.

I. Introduction

The class imbalance problem occurs when the classification results show that there is a class with a much larger number of instances than the other classes. This raises the issue of majority versus minority [1]. Most of the traditional machine learning techniques deal with reasonably balanced class distributions in solving classification problems [2]. The problem of class imbalance is a major problem in machine learning because it can affect predictive accuracy, where machine learning will provide better accuracy of predicted results for classes with larger instances of instances and the resulting accuracy results will be lower in the class with small number of instances [3]. This problem also increase the number of missclassification in machine learning [4]. In General, there are three types of approaches that can be used to solve class imbalance problems. The first approach is Data-Level Solutions, the second approach is Algorithm-Level Solutions, and the third approach is Cost-Sensitive Solutions [5]. The problem of class imbalance can be divided into two scenarios: binary and multi-class [6].

On completion with each approach it has been suggested that the representativeness of the sample problem is important to determine because it involves three things: it must have a reduced size compared to the original source, containing the main information in source, and low redundancy [7]. SVM is robust and precise, but can be sensitive to missing values and difficult to train for large scale data and highly imbalance datasets. For train large scale data, it must be combined with another approach in Data Level-Solutions [8]. A widely used method is Synthetic Minority Over-sampling TEchnique (SMOTE) method [9]. For controlling the sensitivity in Support Vector Machine, [10] has been introduced Biased Support Vector Machine. Biased Support Vector Machine was combined with SMOTE by [11] in handling class imbalance problem.

The SMOTE method will generally initialise a sample number of minorities and after that it will generate equal numbers of synthetic data, but this method should also respect to all the remaining minority data samples [12]. A Weighted-SMOTE proposed by [12] will oversampling of each minority data sample is carried out based on the weight assigned to it. These weights are determined by using the Euclidean distance of a particular minority data sample with respect to all the remaining minority data samples.

This paper introduced the implementation of Biased Support Vector Machine and Weighted-SMOTE in Handling Class Imbalance Problem. Biased Support Vector Machine will give the better sensitivity control compare to Support Vector Machine. Biased Support Vector Machine works by providing different cost functions for positive samples (minority class) and negative samples (majority class) will group both positive samples and negative samples into SV Sets and NSV Sets.

Biased Support Vector Machine has been combined with SMOTEBoost in handling class imbalance problem to get the better accuracy. SMOTEBoost should also respect to all the minority data samples and for this reason, Weighted-SMOTE has been introduced. NSV Sets from negative samples and SV Sets from Positive Samples will then undergo a Weighted-SMOTE process. The traditional performance evaluation in imbalanced evaluation are accuracy and sensitivity [3].

The main contribution in this paper is that once the case if the focus would be placed on the majority class, one of the Method is Biased Support Vector Machine (BSVM) and how to get the better sensitivity and accuracy using BSVM with the implementation of BSVM and Weighted-SMOTE in handling class imbalance problem.

Comment [Office1]: The structure of the presentation should be improved to show the contribution of the paper.

II. Biased Support Vector Machine and Weighted-SMOTE Algorithm

Traditional Support Vector Machine (SVM) cannot handling the class imbalance problem with the bias in the data distribution is significant, it is because the separation hyperplane learned by the SVM is very close to the minority class which explains why SVM has a degrading performance on highly imbalanced datasets [13]. A Geometric Mean SVM (GSVM) proposed by [14] will get the ideal hyperplane by calculating the moving of the original bias in the SVM to improve the geometric mean between the true positive rate and the true negative rate. But in the some problem, The focus of interest is given more to the positive class since it may contain unusual behaviour differs from the general access pattern and to deal with datasets where one class, the positive class, is considered more relevant than another class in binary classification problems, the Biased Support Vector Machine (BSVM) method [10] can be considered.

BSVM is designed for the case when it is non-critical to increase the true positive ratio in exchange for an increase in the false positive rate[15]. BSVM Method will achieve better performance in sensitivity, but have a significant reduction in accuracy[15] and this problem will overcome using the integration with Weighted-Smote Algorithm. Biased Support Vector Machine has been combined with SMOTEBoost in handling class imbalance problem[11] but this method should also respect to all the remaining minority data samples [12] and this is the reason why Weighted-SMOTE has and advantage to use in this reasearch.

In Biased Support Vector Machine, The minority samples are given with a larger cost function and the mathematical model of Biased Support Vector Machine is defined by [10]

Comment [Office2]: In order to show the advantages of the proposed methods, the authors should compare the existing or common methods versus the proposed methods. Analyze the results and show the improvements.

$$min\frac{1}{2}|w|^2 + C^+\sum_{i\in I_+}\xi_i^k + C^-\sum_{i\in I_-}\xi_i^k$$
$$s.t. y_i(w^T\phi(x_i)+b) \geq 1-\xi_i \quad \xi_i \geq 0, \forall i \quad (1)$$

Where w and b represent the normal vector and the intercept of hyperplane, respectively $\xi_i \geq 0$ are slack variable. The detailed procedure of Biased Support Vector Machine and Weighted-SMOTE Algorithm in handling class imbalance is as follows.

Step 1, Take some data from minority class and majority class. Some data from the minority class are expressed as positive samples and some data from the majority class is expressed as negative samples.

Step 2, Determine the hyperplane combination of positive samples and negative samples. Where the combined value of positive samples is 1 and the combined value of negative samples is -1.

Step 3, Hyperplane is obtained by using Biased Support Vector Machine which is done by minimizing the margin value of positive samples and negative samples using equation (2) and (3) [11]

$$\frac{1}{2}\|W\|^2 = \frac{1}{2}(w_1^2 + w_2^2 + w_3^2 + w_4^2) \tag{2}$$

s.t.

$$y_i(w_i x_i + b) \geq 1, \qquad i = 1,2,3,\cdots,N \tag{3}$$

$$y_i(w_1.x_1 + w_2.x_2 + w_3.x_3 + w_4.x_4 + b) \geq 1$$

Step 4, Equation (4) is an equation for the determination of Positive Samples and Negative Samples [11]

$$f(x) = sign\left(\sum_{i=1}^{l} y_i(w.x_i) + b\right) \tag{4}$$

If the sign function gives a result greater than 0 then it is included in the minority class and if the sign function gives a result smaller than 0 then it is included in the majority class.

Conduct training for minority class and majority class based on hyperplane equation obtained. Training process can be done using (4)

Step 5, The result for minority class must be positive and the result for majority class must be negative. If the training result for an instance in the minority class gives a negative result then move the instance into the majority class and if the training result for an instance in the majority class gives a positive result then move that instance into the minority class.

Step 6, Take some data back from the minority class and make it as positive samples and retrieve some data back from majority class and make it as negative samples. Determine the hyperplane of each positive samples and negative samples using (2) and (3).

Step 7, Perform training process on minority class by using (5). for the process of determining SV Sets and NSV Sets by using hyperplane from positive samples. If the calculation result gives then it is categorized into SV Sets and otherwise categorized as NSV Sets.

$$H_1 = SV\ Sets\ for\ Minority\ Class\ (w.x\ +\ b = 1) \tag{5}$$

Step 8, Do the training process on majority class by using (6). for the process of determining SV Sets and NSV Sets by using hyperplane of negative samples. If the calculation result gives then it is categorized into SV Sets and otherwise categorized as NSV Sets.

$$H_2 = SV\ Sets\ for\ Majority\ Class\ (w.x\ +\ b = -1) \tag{6}$$

Step 9, SV Sets on the Minority Class will be eliminated noise and then will process the Weighted-SMOTE process to then be combined with NSV Sets on the minority class to become a new minority class. The Weighted-SMOTE process is as follows [12]

(1) Calculate euclidean distance for the each of the *T* minority data samples using (7) [16]

$$d_{ij} = \sqrt{\sum_{k=1}^{p}(x_{ik} - x_{jk})^2} \tag{7}$$

Here, $i = [1,2,\ldots.,T]$ and $j = [1,2,\ldots,T]$ and $j \neq 1$. The *Euclidean Distance* for all the minority data are calculated and stored in a column matrix using (8) [10]

$$ED = [ED_1, ED_2, \ldots, ED_T] \tag{8}$$

(2) Normalized the ED matrix using the maximum of the ED (ED_{max}) and the minimum of ED (ED_{min}) and called it as normalized ED (NED) that can be seen in (9).

$$NED_i = \frac{ED_i - ED_{min}}{ED_{max} - ED_{min}} \tag{9}$$

(3) Convert the *NED* matrix using the (10) to get the Remodeled Normalized Euclidean Distance Matrix (*RNED*)

$$[RNED]_{Tx1} = sum(NED) - [NED]_{Tx1} \tag{10}$$

(4) Calculated the Weight Matrix for each minority of *T* samples with respect to the sum of RNED Matrix using (11)

$$[Weight\ Matrix]_{Tx1} = \frac{[RNED]_{Tx1}}{sum(RNED)} \tag{11}$$

(5) Using this Weight matrix to find the SMOTE generation matrix using (12)

$$[SMOTE\ Generation\ Matrix]_{Tx1} = \frac{NxT}{100}[Weight\ Matix]_{Tx1} \tag{12}$$

Step 10, NSV Sets on Majority Class will experience Weight-SMOTE process and will then be combined with SV Sets on Majority Class to become the new majority class. The Weighted-SMOTE process is same with *Step* 9.

III. Method

The experiments are conducted using *Iris Dataset* and *Balanced Scale Weight & Distance Database* and the software is based on R Language. The Research Method can be seen in Fig. 1

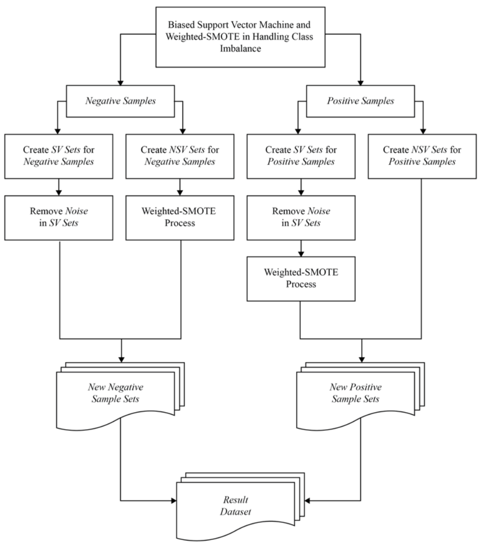

Fig. 1. Research Method

From the Fig. 1, it can be seen that the Biased Support Vector Machine will classify both majority class (negative samples) and minority class (positive samples) into Support Vector Sets (SV Sets) and Non Support Vector Sets (NSV Sets). The noise from the SV Sets Negative Samples will be removed and the NSV Sets Negative Samples will be processed using Weighted-SMOTE. Then SV Sets Negative Samples will combine with NSV Sets Negative Samples, will make a new Negative Samples Sets. After this process, the Noise from the SV Sets in Positive Samples will be removed and this SV Sets will be processed using Weighted-SMOTE. Then SV Sets Positive Samples will combine with NSV Sets Positive Samples, will make a new Positive Samples Sets. New Negative Sample Sets and New Positive Sample Sets will combine to make Result Dataset.

This research will calculate the Accuracy and Sensitivity from the result of implementation of Biased Support Vector Machine and Weighted-SMOTE using (13) [12] and (14) [10].

$$Accuracy = \frac{TN+TP}{TN+FP+FN+TP} \tag{13}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{14}$$

Using Confusion Matrix [10] for determine the value of TN, TP, FN, and FP that can be seen inf Table 1.

Table 1. Confusion Matrix

	Classified as Positive	Classified as Negative
Positive Samples	*True Positive* (TP)	*False Negative* (FN)
Negative Samples	*False Positive* (FP)	*True Negative* (TN)

IV. Results and Discussion

The testing are conducted using *Iris Dataset* and *Balanced Scale Weight & Distance Database* and the software is based on R Language. The result of the testing using *Iris Dataset* can be seen in Table 2.

Table 2. The Result Testing using *Iris Dataset*

Iteration	Before Processing		After Processing		Accuracy	Sensitivity
	Number of Majority	Number of Minority	Number of Majority	Number of Minority		
1	60	40	52	48	0.87	0.83
2	60	34	49	45	0.81	0.79
3	62	39	52	49	0.82	0.79
4	64	36	50	50	0.902	0.904
5	55	44	50	49	0.93	0.92
6	71	29	50	50	0.84	0.83
7	66	30	50	46	0.93	0.92
8	62	37	50	49	0.89	0.88
9	67	35	54	48	0.94	0.93
10	65	35	52	48	0.89	0.88
Average	63.2	35.9	50.9	48.2	0.88	0.87

The result of the testing using *Balanced Scale Weight & Distance Database Dataset* can be seen in Table 3.

Table 3. The Result Testing using *Balanced Scale Weight & Distance Database*

Iteration	Before Processing		After Processing		Accuracy	Sensitivity
	Number of Majority	Number of Minority	Number of Majority	Number of Minority		
1	325	251	290	286	0.83	0.832
2	353	226	291	288	0.84	0.82
3	345	242	301	286	0.91	0.906
4	363	213	295	281	0.87	0.85
5	347	232	291	288	0.83	0.84
6	319	257	292	284	0.88	0.87
7	321	256	294	283	0.9	0.91
8	316	260	293	283	0.86	0.85
9	343	236	290	289	0.79	0.8
10	361	219	294	286	0.85	0.84
Average	339.3	239.2	293.1	285.4	0.86	0.85

From the result that can be seen in Table 2 dan Table 3, there are an improvement in the different of number of majority and number of minority before and after processing with Biased Support Vector Machine and Weighted-SMOTE. The accuracy and sensitivity that has been achieved from the testing is very satisfied. The high value of Accuracy and Sensitivity means that the implementation of Biased Support Vector Machine and Weighted-SMOTE can handle the class imbalance problem.

V. Conclusion

This paper has introduced the implementation of Biased Support Vector Machine and Weighted-SMOTE in handling class imbalance problem. The implementation is conducted using Iris Dataset and Balanced Scale Weight & Distance Database. The result of this testing has given the very satisfied result that indicate in value of the Accuracy and Sensitivity. for future research should increase the number of datasets involved in the test so that the results obtained can be more accurate.

References

[1] S. M. A. Elrahman and A. Abraham, “A Review of Class Imbalance Problem,” *J. Netw. Innov. Comput.*, vol. 1, pp. 332–340, 2013.
[2] J. Ortigosa-Hernández, I. Inza, and J. A. Lozano, “Measuring the class-imbalance extent of multi-class problems,” *Pattern Recognit. Lett.*, vol. 98, pp. 32–38, 2017.
[3] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, “A Review on Ensembles for the Class Imbalance Problem: Bagging-, Boosting-, and Hybrid-Based Approaches,” *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 42, no. 4, pp. 463–484, Jul. 2012.
[4] Hartono, O. S. Sitompul, Tulus, and E. B. Nababan, “Optimization Model of K-Means Clustering Using Artificial Neural Networks to Handle Class Imbalance Problem,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 288, p. 012075, Jan. 2018.
[5] W.-C. Lin, C.-F. Tsai, Y.-H. Hu, and J.-S. Jhang, “Clustering-Based Undersampling in Class-Imbalanced Data,” *Inf. Sci.*, vol. 409–410, pp. 17–26, May 2017.
[6] J. A. Sáez, B. Krawczyk, and M. Woźniak, “Analyzing the oversampling of different classes and types of examples in multi-class imbalanced datasets,” *Pattern Recognit.*, vol. 57, pp. 164–178, Sep. 2016.
[7] D. Furundzic, S. Stankovic, S. Jovicic, S. Punisic, and M. Subotic, “Distance based resampling of imbalanced classes: With an application example of speech quality assessment,” *Eng. Appl. Artif. Intell.*, vol. 64, pp. 440–461, Sep. 2017.

[8] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, and G. Bing, “Learning from Class-Imbalanced Data: Review of Methods and Applications,” *Expert Syst. Appl.*, vol. 73, pp. 220–239, May 2017.
[9] L. Zhang, C. Zhang, R. Gao, R. Yang, and Q. Song, “Using the SMOTE technique and hybrid features to predict the types of ion channel-targeted conotoxins,” *J. Theor. Biol.*, vol. 403, pp. 75–84, 2016.
[10] K. Veropoulos, C. Campbell, and N. Cristianini, “Controlling the Sensitivity of Support Vector Machines,” presented at the International Joint Conference on AI, 1999, pp. 55–60.
[11] C. Jian, J. Gao, and Y. Ao, “A new sampling method for classifying imbalanced data based on support vector machine ensemble,” *Neurocomputing*, vol. 193, pp. 115–122, Jun. 2016.
[12] M. R. Prusty, T. Jayanthi, and K. Velusamy, “Weighted-SMOTE: A Modification to SMOTE for Event Classification in Sodium Cooled Fast Reactors,” *Prog. Nucl. Energy*, vol. 100, pp. 355–364, 2017.
[13] H. He and A. Ghodsi, “Rare class classification by support vector machine,” in *Pattern Recognition (ICPR), 2010 20th International Conference on*, 2010, pp. 548–551.
[14] L. Gonzalez-Abril, H. Nuñez, C. Angulo, and F. Velasco, “GSVM: An SVM for handling imbalanced accuracy between classes inbi-classification problems,” *Appl. Soft Comput.*, vol. 17, pp. 23–31, Apr. 2014.
[15] L. Gonzalez-Abril, C. Angulo, H. Nuñez, and Y. Leal, “Handling binary classification problems with a priority class by using Support Vector Machines,” *Appl. Soft Comput.*, vol. 61, pp. 661–669, Dec. 2017.
[16] R. Loohach and K. Garg, “Effect of Distance Functions on K-Means Clustering Algorithm,” *Int. J. Comput. Appl.*, vol. 49, no. 6, pp. 7–9, 2012.

International Journal of Advances in Intelligent Informatics
ISSN 2442-6571 (print) 2548-3161 (online) | available at http://ijain.org/index.php/IJAIN
Member of Crossref
Accreditated "1st" grade (SINTA 1) by Ministry of Research, Technology and Higher Education (RistekDikti) of the Republic of Indonesia, Decree No. 30/E/KPT/2018, October 24, 2018
Indexed in Scopus, sinta, ASEAN Citation Index, EBSCOhost, Academic Search, ProQuest
Published by

#146 Editing

SUMMARY REVIEW **EDITING**

Submission

Authors	Hartono Hartono, Opim Salim Sitompul, Tulus Tulus, Erna Budhiarti Nababan
Title	Biased support vector machine and weighted-smote in handling class imbalance problem
Section	Articles
Editor	Mohammad Syafrullah

Copyediting

COPYEDIT INSTRUCTIONS

Copyeditor Dr. Aji Prasetya Wibawa

REVIEW METADATA		REQUEST	UNDERWAY	COMPLETE
1.	Initial Copyedit File: None	2018-03-13	—	—
2.	Author Copyedit File: None Choose File No file chosen Upload	—	—	
3.	Final Copyedit File: None	—	—	—

Copyedit Comments No Comments

Layout

Layout Editor None

QUICK MENU
Editorial Team
Reviewer
Focus and Scope
Author Guidelines
Publication Ethics
Open Access Policy
Peer Review Process
Online Submission
Author(s) Fee
Contact

USER
You are logged in as...
hartono
- My Profile
- Log Out

TEMPLATE

Layout Version	REQUEST	UNDERWAY	COMPLETE	VIEWS
146-628-1-LE.DOCX 2018-05-13	—	—	—	

Galley Format	FILE		
1. PDF VIEW PROOF	146-629-3-PB.PDF 2018-05-14		482

Supplementary Files	FILE
	None

Layout Comments No Comments

Proofreading

Proofreader None

REVIEW METADATA

	REQUEST	UNDERWAY	COMPLETE
1. Author	—	—	
2. Proofreader	—	—	—
3. Layout Editor	—	—	—

Proofreading Corrections No Comments PROOFING INSTRUCTIONS

International Journal of Advances in Intelligent Informatics
ISSN 2442-6571 (print) | 2548-3161 (online)
Organized by UAD and ASCEE Computer Society
Published by Universitas Ahmad Dahlan
W: http://ijain.org
E: info@ijain.org (paper handling issues)
andri.pranolo.id@ieee.org (publication issues)

Academic Search

Crossref

EBSCOhost

Google Scholar

ProQuest

More Abstracting & Indexing...

REQUEST INDEXING

- **SCOPUS (ACCEPTED)**
  - Submission Received: **July 26, 2017**
  - Submission Accepted: **June 5, 2018**
  - **SCOPUS CiteScore Tracker 2020**
- **Web of Science**
  - Latest submission: June 11, 2018
  - **Web of Science** Citation Analysis
- **IET INSPEC**
  - Added to review: **June 5, 2018**
- **COMPENDEX**
  - Submission: **October 15, 2017**
  - Resubmission: **September 27, 2020**

CURRENT ISSUE

ATOM 1.0
RSS 2.0
RSS 1.0

KEYWORDS

CNN COVID-19 Class Imbalance Classification Clustering Computer vision Deep Learning Deep learning Effort estimation Feature selection GPU LSTM Machine Learning Machine learning Neural network Ontology Optimization SVM Scheduling Support Vector Machines (SVM) Transfer Learning

Editor Decision Accepted

[IJAIN] Editor Decision

Kotak Masuk ×

Andri Pranolo <andri.pranolo@tif.uad.ac.id>
kepada saya

Sel, 13 Mar 2018, 11.29

Mr. Hartono -:

We have reached a decision regarding your submission to International Journal of Advances in Intelligent Informatics, "Implementation of Biased Support Vector Machine and Weighted-SMOTE in Handling Class Imbalance Problem".

Our decision is to: Accept Submission

Please keep attention for the copy editing and proofreading process which are final publicity process on IJAIN Journal. Your paper is scheduled to be published in the upcoming Vol. 4, No. 1, March 2018 issue after we finished those process.

Andri Pranolo
(Managing Editor)

International Journal of Advances in Intelligent Informatics (IJAIN)
ISSN 2442-6571 (print) | 2548-3161 (online)
http://ijain.org/index.php/IJAIN
Email: ijain@uad.ac.id, andri.pranolo@tif.uad.ac.id

Bukti Copyediting

#146 Editing

SUMMARY REVIEW **EDITING**

Submission

Authors	Hartono Hartono, Opim Salim Sitompul, Tulus Tulus, Erna Budhiarti Nababan
Title	Biased support vector machine and weighted-smote in handling class imbalance problem
Section	Articles
Editor	Mohammad Syafrullah

Copyediting

COPYEDIT INSTRUCTIONS

Copyeditor Dr. Aji Prasetya Wibawa

REVIEW METADATA		REQUEST	UNDERWAY	COMPLETE
1.	Initial Copyedit File: None	2018-03-13	—	—
2.	Author Copyedit File: None Choose File No file chosen Upload	—	—	
3.	Final Copyedit File: None	—	—	—

Copyedit Comments No Comments

Layout

Layout Editor None

Layout Version	REQUEST	UNDERWAY	COMPLETE	VIEWS
146-628-1-LE.DOCX 2018-05-13	—	—	—	

Galley Format	FILE	VIEWS
1. PDF VIEW PROOF	146-629-3-PB.PDF 2018-05-14	482

Supplementary Files	FILE