HAR-MI method for multi-class imbalanced datasets

Hartono
Universitas Potensi Utama

Yeni Risyani
STMIK IBBI

Erianto Ongko
Akademi Teknologi Industri Immanuel

Dahlan Abdullah
Universitas Malikussaleh

TELKOMNIKA: Telecommunication Computing Electronics and Control Vol. 18 No. 2, April 2020

1. First Submit : 15 Juni 2019
2. Initial Screening Phase: 21 Juli 2019
3. Editor Decision Accepte with Minor Revision : 12 September 2019
4. Registrasi untuk ICW Telkomnika 2019: 28 September 2019
5. Registrasi Diterima: 7 Oktober 2019
6. Editor Decision Accepted for Telkomnika: 22 Desember 2019
7. Proof Reading: 19 Februari 2020
8. Terbit: April 2020

Bukti Korespondensi Pengajuan Guru Besar Hartono

Judul Karya Ilmiah	:	HAR-MI method for multi-class imbalanced datasets
Jurnal	:	TELKOMNIKA: Telecommunication Computing Electronics and Control)
Penulis	:	Hartono, Yeni Risyani, Erianto Ongko, Dahlan Abdullah
Edisi	:	2020
Volume	:	Vol. 18 No. 2
Penerbit	:	Universitas Ahmad Dahlan

UAD Universitas Ahmad Dahlan

TELKOMNIKA (Telecommunication Computing Electronics and Control)

http://telkomnika.uad.ac.id/index.php/TELKOMNIKA
ISSN 1693-6930, e-ISSN 2302-9293

CiteScore 2021 3.8 Scopus | Q3 | SJR 2022 0.286 | SNIP 2022 0.558

HOME ABOUT LOGIN REGISTER SEARCH CURRENT ARCHIVES ANNOUNCEMENTS

Home > Vol 18, No 2 > **Hartono**

HAR-MI method for multi-class imbalanced datasets

H. Hartono, Yeni Risyani, Erianto Ongko, Dahlan Abdullah

Abstract

Research on multi-class imbalance from a number of researchers faces obstacles in the form of poor data diversity and a large number of classifiers. The Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) method is a Hybrid Ensembles method which is the development of the Hybrid Approach Redefinion (HAR) method. This study has compared the results obtained with the Dynamic Ensemble Selection-Multiclass Imbalance (DES-MI) method in handling multiclass imbalance. In the HAR-MI Method, the preprocessing stage was carried out using the random balance ensembles method and dynamic ensemble selection to produce a candidate ensemble and the processing stages was carried out using different contribution sampling and dynamic ensemble selection to produce a candidate ensemble. This research has been conducted by using multi-class imbalance datasets sourced from the KEEL Repository. The results show that the HAR-MI method can overcome multi-class imbalance with better data diversity, smaller number of classifiers, and better classifier performance compared to a DES-MI method. These results were tested with a Wilcoxon signed-rank statistical test which showed that the superiority of the HAR-MI method with respect to DES-MI method.

Keywords

classifier; data diversity; hybrid approach redefinition-multiclass imbalance; multi-class imbalance;

Full Text:

PDF

DOI: http://doi.org/10.12928/telkomnika.v18i2.14818

USER

Username
Password
Remember me
Login

QUICK LINKS

- Editorial Boards
- Reviewers
- **Author Guidelines**
- **Online Submission**
- Policy of Plagiarism Screening
- **Peer Review Process**
- **Publication Fee**
- Abstracting and Indexing
- Repository Policy
- Scopus: Add missing document
- Publication Ethics
- Visitor Statistics
- Contact Us

Q2 CiteScore Quartile in Electrical and Electronic Engineering

3.8 CiteScore 2022

56th percentile

Powered by Scopus

JOURNAL CONTENT

Keterangan Bahwa semua paper accepted padaa ICW Telkomnika akan Dipublikasikan pada Jurnal Telkomnika ataupun Jurnal lain sesuai scope

Gmail Hartono Ibbi <hartonoibbi@gmail.com>

[Scopus-indexed] Paper submission deadline: Jul 20, 2019 (Extended)

1 pesan

Dr. Tole Sutikno <telkomnika@uad.ac.id> 11 Juli 2019 pukul 05.47
Kepada: Hartono Hartono <hartonoibbi@gmail.com>

2019 2nd International Conference and Workshop on Telecommunication, Computing, Electronics and Control (ICW TELKOMNIKA 2019)
Call for Papers [Apologies if you receive multiple copies of this announcement]

Dear Colleague,

Please consider to submit your papers for the 2019 2nd ICW-TELKOMNIKA. This event will be held on November 19-22, 2019 at Royal Ambarrukmo Hotel, Yogyakarta, Indonesia. This conference is hosted by Universitas Ahmad Dahlan and organized by TELKOMNIKA Telecommunication Computing Electronics and Control (Scopus indexed journal).

All presented papers in this conference (after extension & revision based upon feedback at the conference & workshop) will be published in one of the journals below:
- TELKOMNIKA Telecommunication Computing Electronics and Control (Scopus indexed journal, SJR Q2)
- Bulletin of Electrical Engineering and Informatics (Scopus indexed journal, SJR Q2)
- International Journal of Power Electronics and Drive Systems (Scopus indexed journal, SJR Q2).

IMPORTANT DATES
Conference Date : Nov 19-22, 2019

Batch II
Submission deadline : Jul 20, 2019 (Extended)
Notification of Acceptance : Aug 30, 2019
Camera ready-papers deadline : Oct 10, 2019
Registration deadline : Oct 25, 2019

Please distribute this call for papers among your colleagues, PhD students, and young researchers.

Thanks for the continuing interest in our works and events.

Best Regards,
Dr. Tole Sutikno
General Chair
2019 2nd ICW-TELKOMNIKA Int Conf
Website: http://ojs.telkomnika.com/index.php/ICWT
email: telkomnika@uad.ac.id, icwt1@ee.uad.ac.id

2nd 2019 International Conference and Workshop on Telecommunication, Computing, Electronics and Control (ICW-TELKOMNIKA 2019)
http://ojs.telkomnika.com/index.php/ICWT

First Submit

Original File in Submission Not Uploaded (ID Submission: 53, Section: Computing)

1 pesan

Hartono Ibbi <hartonoibbi@gmail.com> 15 Juni 2019 pukul 13.46
Kepada: Tole Sutikno <tole@ee.uad.ac.id>

Dear Mr. Tole Sutikno,

I have submitted a paper for participation in ICW-Telkomnika with ID: 53 for the Computing category. But, apologize in advance if I didn't press the upload button so I suspect this time the original file in my submission is empty. Please inform me, if I am allowed to send my submission file via e-mail or can I submit it via supplementary files or I make a new submission or maybe there are other things that you recommend.

Thank you for your help.

Best Regards,

Hartono

Initial Screening Phase

[ICWT] [Editor] An initial screening phase "Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) Method for multi-class imbalanced datasets"

3 pesan

Dr. Tole Sutikno <telkomnika@uad.ac.id> 10 Juli 2019 pukul 23.09
Kepada: Hartono Hartono <hartonoibbi@gmail.com>

Dear Prof/Dr/Mr/Mrs: Hartono Hartono,

We have reached an initial screening phase regarding your paper submission entitled "Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) Method for multi-class imbalanced datasets" to ICW TELKOMNIKA.

Our decision is: Revisions Required.
A high quality paper should has:
(1) a clear statement of the problem the paper is addressing;
(2) the proposed solution(s); and
(3) results achieved. It describes clearly what has been done before on the problem, and what is NEW.

The goal of your first revisions is to describe NOVEL technical results.
Please note that there are four (4) types of technical results:
(1) An algorithm;
(2) A system construct: such as hardware design, software system, protocol, etc.;
(3) A performance evaluation: obtained through analyses, simulation or measurements;
(4) A theory: consisting of a collection of theorems.

Your revisions should focus on:
(1) Describing the results in sufficient details to establish their validity;
(2) Identifying the novel aspects of the results, i.e., what new knowledge is reported and what makes it non-obvious;
(3) Identifying the significance of the results: what improvements and impact do they suggest.

The main goal of this stage is to ensure that the next person who designs a system like yours doesn't make the same mistakes and takes advantage of some of your best solutions. So make sure that the hard problems (and their solutions) are discussed and the non-obvious mistakes (and how to avoid them) are discussed.

Read the checklist for preparing your revisions at:
http://journal.uad.ac.id/index.php/TELKOMNIKA/about/editorialPolicies#custom-1.
Please try to follow the format and guide (http://goo.gl/FiPFbF) as closely as possible. Original/Research paper should be presented with IMRaD style/model:
1. Introduction
2. The Proposed Method/Algorithm/Procedure specifically designed (optional). Authors may present complex proofs of theorems or non-obvious proofs of correctness of algorithms after introduction section (obvious theorems & straightforward proofs of existing theorems are NOT needed).
3. Research Method
4. Results and Discussion
5. Conclusion.

We will usually expect a minimum of 20 to 25 references primarily to journal papers, depending on the length of the paper. You can found our published papers to enrich your references at:
- http://journal.uad.ac.id/index.php/telkomnika

- http://beei.org
- http://ijpeds.iaescore.com

Submit your revised paper within 12 days, and do serious work for updating your paper. Revised paper submission is submitted (as author version) on the same paper ID number through our online system. When your revised paper reached us, it will be sent for single blind peer review by at least two reviewers who will either be members of the Editorial Board or others of similar standing in the field, for contribution, originality, relevance, and presentation. Then, your paper will be judged for final decision of acceptance or rejection.

All presented papers in this conference (after extension & revision based upon feedback at the conference & workshop) will be published in one of the journals below:
- TELKOMNIKA Telecommunication Computing Electronics and Control (Scopus indexed journal, SJR Q2)
- Bulletin of Electrical Engineering and Informatics (Scopus indexed journal, SJR Q2)
- International Journal of Power Electronics and Drive Systems (Scopus indexed journal, SJR Q2).

We look forward to receiving the revised version of your manuscript and are delighted that you chose to send this important work to this event.

Best Regards,
Dr. Tole Sutikno
General Chair
2019 2nd International Conference and Workshop on Telecommunication, Computing, Electrical, Electronics and Control (ICW-TELKOMNIKA 2019)
telkomnika@uad.ac.id
Website: http://ojs.telkomnika.com/index.php/ICWT
email: icwt1@ee.uad.ac.id

2nd 2019 International Conference and Workshop on Telecommunication, Computing, Electronics and Control (ICW-TELKOMNIKA 2019)
http://ojs.telkomnika.com/index.php/ICWT

Hartono Ibbi <hartonoibbi@gmail.com> 21 Juli 2019 pukul 21.29
Kepada: "Dr. Tole Sutikno" <telkomnika@uad.ac.id>

Dear Dr. Tole Sutikno
General Chair
2019 2nd International Conference and Workshop on Telecommunication, Computing, Electrical, Electronics and Control (ICW-TELKOMNIKA 2019)

We have completed the revision of our paper in accordance with the input given by the Reviewer and have uploaded our revised paper by OJS. Thank you for this very meaningful input. We realize that there is still a possibility that the revision that we have done is not fully in accordance with what the Reviewer team wants. On this occasion, we hope that the reviewer team will not hesitate to tell us if there is something that needs to be improved. Thank you if We are allowed to take part in the ICW-TELKOMNIKA 2019.

Best Regards,

Hartono
[Kutipan teks disembunyikan]

Jurnal TELKOMNIKA <telkomnika@uad.ac.id> 21 Juli 2019 pukul 21.29
Kepada: hartonoibbi@gmail.com

Dear Sir/Madam,

Your email has reached my queues safely, so i will get back in touch with you there as soon as i can. Just a heads up i might not get back to you straight away, as i answer all emails in order to keep things fair.

How to submit your manuscript

To make a submission, you must have a user account and be enrolled as an Author. User accounts can either be created by the Journal Manager or you can register yourself (this journal policy allow you create user account by your self as a Reader, an Author and/or a Reviewer). All fields with an asterisk beside them (Username; Password; Repeat Password; First Name; Last Name; Email) are mandatory. Your username and your email address must be unique; furthermore, while you can change your email address at a later date, you will be unable to change your username. If you want to register in another role within the same journal (for example, if you are already a Reader, but also want to become an Author) you can log in; go to Edit My Profile (under My Account on your User Home page); and check off the checkboxes next to any available roles, near the bottom of the page. Once you have an account, log in to the journal site and select the role of Author. The Author is asked to upload a submission file and to provide metadata or indexing information. (The metadata improves the search capacity for research online and for the journal.) The Author can upload Supplementary Files, in the form of data sets, research instruments, or source texts that will enrich the item, as well as contribute to more open and robust forms of research and scholarship.

Your paper ID?
Please ALWAYS put your paper ID number in the subject line of email communication when making inquiries. For example, if your submitted manuscript URL: http://journal.uad.ac.id/index.php/TELKOMNIKA/author/submission/1234, your paper ID number is 1234.

Your cooperation is very appreciated.

Thank you

Best Regards,
Tole Sutikno
Editor:
TELKOMNIKA TCEC, http://journal.uad.ac.id/index.php/TELKOMNIKA

Journal Form for Hardcopy Order

Please download "Journal form for Hardcopy order <http://goo.gl/a6Y7zg>" and contact us if further information is needed:

Mr. Sulchan Hidayat: sulchan.hidayat@gmail.com, +62 81328731075

••••••••••••••••••••
Our event: An International Conference Indexed by IEEEXplore and Scopus
"4th IEEE Conference on Energy Conversion (CENCON 2019)"
Website: http://cencon.intconference.org
EDAS papers submission: https://edas.info/N25880

The 4th IEEE Conference on Energy Conversion (CENCON 2019) will be held at Royal Ambarrukmo hotel, located in the capital of Yogyakarta, Indonesia on 16 - 17 October 2019. The objective of the conference is to share the latest research in the areas of power electronics, electrical drives, renewable energy, modeling and control systems. CENCON 2019 is sponsored by the IEEE Malaysia Power Electronics (PEL) Chapter, co-organized by IEEE Indonesia Power Electronics (PEL) Chapter and Universitas Ahmad Dahlan and technical co-sponsored by The Korean Institute of Power Electronics (KIPE) and Power Electronics and Drives Research Group (PEDG).

--

Best Regards,
Tole Sutikno
Editor-in-Chief, TELKOMNIKA (Telecommunication, Computing, Electronics and Control)
ISSN: 1693-6930, e-ISSN: 2302-9293
http://journal.uad.ac.id/index.php/TELKOMNIKA

UNIVERSITAS AHMAD DAHLAN

Kampus 1: *Jln. Kapas No. 9 Yogyakarta*

Kampus 2: *Jl. Pramuka 42, Sidikan, Umbulharjo, Yogyakarta 55161*

Kampus 3: *Jl. Prof. Dr. Soepomo, S.H., Janturan, Warungboto, Umbulharjo, Yogyakarta 55164*

Kampus 4: *Jl.Ringroad Selatan, Yogyakarta*

Kampus 5: *Jl. Ki Ageng Pemanahan 19, Yogyakarta*

Kontak

Email: info@uad.ac.id

Telp. : (0274) 563515, 511830, 379418, 371120

Fax. : (0274) 564604

Decision Accepted with Minor Revision

Gmail

Hartono Ibbi <hartonoibbi@gmail.com>

[ICWT] #53 Editor Decision "Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) Method for multi-class imbalanced datasets"

2 pesan

IAES CFP <cfp@iaescore.com> 12 September 2019 pukul 07.22
Kepada: hartonoibbi@gmail.com
Cc: icwt1@ee.uad.ac.id, lppi@uad.ac.id

Editor
2019-09-09 06:23 AM

Dear Prof/Dr/Mr/Mrs: Hartono Hartono,

We have reached a decision regarding your submission entitled "Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) Method for multi-class imbalanced datasets" to ICW TELKOMNIKA has been ACCEPTED with minor revisions. Congratulations!
You are invited to present your paper in this international conference and workshop.

Please prepare your final camera ready paper (in MS Word file format) adheres every detail of the guide of authors (http://goo.gl/FiPFbF), and check it for spelling/grammatical mistakes. We will usually expect a minimum 25 references primarily to journal papers, depending on the length of the paper.

The goal of this camera ready paper is to describe NOVEL TECHNICAL RESULTS. There are four (4) types of technical results:
1. A theory: consisting of a collection of theorems.
2. An algorithm/method/approach/framework/ .... ;
3. A system construct: such as hardware design, software system, protocol, etc.; The main goal of your revised paper is to ensure that the next person who designs a system like yours doesn't make the same mistakes and takes advantage of some of your best solutions. So make sure that the hard problems (and their solutions) are discussed and the non-obvious mistakes (and how to avoid them) are discussed.
4. A performance evaluation: obtained through analyses, simulation or measurements;
Please state your technical result (one of the four types of technical results above) in your paper.

In order to cover part of the event cost, each accepted paper will be charged as available information at http://icw.telkomnika.com (please refer our flyer).

The payment should be made by bank transfer (T/T):

Bank Account name (please be exact)/Beneficiary: TOLE SUTIKNO
Bank Name: Bank Central Asia (BCA)
Branch Office: Kusumanegara
City: Yogyakarta
Country :Indonesia
Bank Account #: 8465122249
SWIFT Code: CENAIDJAXXX

or through PayPal (as alternative of bank transfer) to email: tole@ee.uad.ac.id
---

Please submit your final camera ready paper and payment receipt within 6 weeks.

I look forward for your response

Sincerely yours,
Dr. Tole Sutikno
General Chair
2019 2nd International Conference and Workshop on Telecommunication, Computing, Electrical, Electronics and Control (ICW-TELKOMNIKA 2019)
telkomnika@uad.ac.id

Bank's detailed address :
Bank BCA Kusumanegara
Jl. Kusumanegara No. 18
City: Yogyakarta
Province: D.I. Yogyakarta (DIY)
Country :Indonesia
Post Code: 55165
Indonesia, Phone:+62 274 418896

The Beneficiary's address:
Kampus 3 Universitas Ahmad Dahlan
Jln. Prof. Soepomo, Janturan
City: Yogyakarta
Province: D.I. Yogyakarta (DIY)
Post Code: 55164
Country: Indonesia

Reviewer J:

Does the paper contain an original contribution to the field?:
No

Is the paper technically sound?:
Yes

Does the title of the paper accurately reflect the major focus contribution of this paper?:
Yes

Please suggest change of the title as appropriate within 12 words:

Is the abstract a clear description of the paper?:
Yes

Please suggest change of the abstract as appropriate within 200 words:

Is the paper well written (clear, concise, and well organized)?:
Yes

Please suggest change for well presentation (clear, concise, and well organized) of this paper

:

Are the equations, figures and tables in this journal style, clear, relevant, and are the captions adequate?:
Yes

Please score the paper on a scale of 1-10 as per the directions below::
6-7 Average

Comments to the Authors:
In this paper, authors propose a novel methodology based on ensembles for the classification of multi-class imbalanced problems. The main idea is to use a combination among preprocessing approaches, bagging, and dynamic classifier selection, in order to boost the recognition ability of the final classification system.

There are some drawbacks that must be addressed prior to the publication of this research:

1. First of all, the methodology itself must be better defined and described in both the abstract and introduction. The term "Hybrid redefinition" is used quite often but the reader is unable to

acknowledge what it is with certainty. Please be more specific.
2. In the abstract, HAR is defined twice (HAR and HAR-MI), whereas DES-MI is not. Please be careful. Also do not use acronyms as keywords.
3. Where are the 20 main research topics listed? A reference is needed
4. When authors point out that there are several approaches to cope with imbalanced, they only mention it but do not describe them. A short definition is needed for the manuscript to be self-content.
5. Revise this and other sentences for typos: "The multi-class imbalance problems that will be solved are problems that involve". Also, Avoid using numbers within the text: "there are 2 things"
6. Divide the introduction section in different paragraphs to ease the reading, especially second paragraph in page 2.
7. In Section 2, it is preferred to explain the method first and then to refer to the pseudo-code. Instead, the algorithm it is very hard to follow in this type of notation. Same for Section 2.2.
8. I do not understand why to divide the content in Section 2 and 3, when these two should be merged.
9. Why it is needed to compute a k-means clustering to measure the imbalance? Since the problem is a supervised one, I do not see the point of this.
10. It is necessary to enlarge the experimental framework to at least 5-10 datasets to ensure a good statistical support for the results.
11. When authors state that they use 10 test, do they refer to a 10 fold cross validation?
12. I guess that the data diversity number in Table 4 refers to the Q-statistic, but this must be made clear. Also, how is the "number of classifiers" set ?It seems too much for this low size problems.
13.The conclusion section should rather be a summary for the whole work. Please use the experimental section to make the whole analysis, and leave the final section for the actual lessons learned and the future work.

2nd 2019 International Conference and Workshop on Telecommunication, Computing, Electronics and Control (ICW-TELKOMNIKA 2019)
http://ojs.telkomnika.com/index.php/ICWT

Hartono Ibbi <hartonoibbi@gmail.com> 12 September 2019 pukul 10.29
Kepada: IAES CFP <cfp@iaescore.com>

Dr. Tole Sutikno
General Chair
2019 2nd International Conference and Workshop on Telecommunication, Computing, Electrical, Electronics and Control (ICW-TELKOMNIKA 2019)

Thank you for the opportunity given to us to participate in the 2019 ICW-TELKOMNIKA. We are also grateful for the valuable input so as to improve the quality of our research. We will revise our writing in accordance with input from reviewers and we will do the payment as soon as possible.

Best Regards,

Hartono
[Kutipan teks disembunyikan]

Registrasi

2019 2nd ICW-TELKOMNIKA Registration Form

1 pesan

Google Forms <forms-receipts-noreply@google.com> 28 September 2019 pukul 13.09
Kepada: hartonoibbi@gmail.com

Thanks for filling out 2019 2nd ICW-TELKOMNIKA Registration Form

Here's what we got from you:

2019 2nd ICW-TELKOMNIKA Registration Form

2019 2nd International Conference and Workshop on Telecommunication, Computing, Electrical, Electronics and Control. November 19-22, 2019 at Royal Ambarrukmo Hotel, Yogyakarta, Indonesia

If you don't have account in gmail, please contact our secretariat.
Secretariat email: icwt1@ee.uad.ac.id, cc: lppi@uad.ac.id

Email address *

hartonoibbi@gmail.com

Paper Submission

Paper ID Number *

53

Title of Paper *

Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) Method for multi-class imbalanced datasets

Tracks *

- ( ) Biomedical Engineering and Bioinformatics
- (•) Computer Science and Applications
- ( ) Electrical Power Engineering

- ( ) Electronics and Instrumentation
- ( ) Information Systems and Technologies
- ( ) Networks and Telecommunication Systems
- ( ) Robotics, Control and Automation
- ( ) Signal, Image and Video Processing
- ( ) Soft Computing and Intelligent System
- ( ) INVITED PAPER

Final Camera Ready Paper *

Files submitted:

Paper ICW Hartono-Blm Final - Hartono Ibbi.docx

Author Information

All Author Names *

Hartono, Erianto Ongko, Dahlan Abdullah

Participate as *

- (•) Presenter
- ( ) Accompanying Person

Participant Full Name *

Hartono

First Affiliation of Participant *

STMIK IBBI

Country of Main Affiliation of Presenter *

Indonesia

Gender *

- (•) Male

Female

Phone *

+62-85270601999

Payment Receipt

Conference (included workshop) Payment Receipt *

Files submitted:

Screen Shot 2019-09-28 at 2.02.23 PM - Hartono Ibbi.png

Create your own Google Form

Gmail

Hartono Ibbi <hartonoibbi@gmail.com>

[ICW-TELKOMNIKA 2019] REGISTRATION SUCCESSFUL!!

2 pesan

icwt1 telkomika <icwt1@ee.uad.ac.id> 7 Oktober 2019 pukul 13.36
Kepada: Hartono Ibbi <hartonoibbi@gmail.com>

Dear Mr./Mrs. H. Hartono,

Your re-registration of paper ID 53 entitled Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) Method for multi-class imbalanced datasets has been accepted in 2nd ICW-TELKOMNIKA 2019. The conference will be held on November 19-21, 2019 at Royal Ambarrukmo Hotel, Yogyakarta, Indonesia. All participants are required to attend all sessions until the end of seminar program. Please plan your trips and lodging on that date. We don't provide the Hotel room for participant lodging. Through this email, we send the payment receipt of conference fee and letter of acceptance. We will inform the schedule of conference as soon as possible.

If you have a question, feel free to contact us. Thank you.

Best regards,
2nd ICW-TELKOMNIKA 2019 Committee

2nd ICW-TELKOMNIKA 2019
International Conference and Workshop on Telecommunication, Computing, Electrical, Electronics and Control

Editor in Chief : telkomnika@uad.ac.id
Secretariat : icwt1@ee.uad.ac.id

Website :
icw.telkomnika.com

13 #53.pdf
141K

Hartono Ibbi <hartonoibbi@gmail.com> 15 Oktober 2019 pukul 12.21
Kepada: icwt1 telkomika <icwt1@ee.uad.ac.id>

Dear 2nd ICW-TELKOMNIKA 2019 Committee ,

Thank you for the assistance given and the opportunity to participate in the 2019 ICW-TELKOMNIKA.

Best Regards,

Hartono
[Kutipan teks disembunyikan]

Gmail Hartono Ibbi <hartonoibbi@gmail.com>

[TELKOMNIKA] Editor Decision "Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) Method for multi-class imbalanced datasets"

5 pesan

Tole Sutikno <telkomnika@uad.ac.id> 22 Desember 2019 pukul 20.05
Balas Ke: "Dr. Tole Sutikno" <telkomnika@uad.ac.id>
Kepada: "Dr. H Hartono" <hartonoibbi@gmail.com>
Cc: Erianto Ongko <eriantoongko@gmail.com>, Dahlan Abdullah <dahlan@unimal.ac.id>, Yeni Risyani <ms_yenir@yahoo.com>

Dear Prof/Dr/Mr/Mrs: Dr. H Hartono,

It is my great pleasure to inform you that your paper entitled "Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) Method for multi-class imbalanced datasets" has been accepted and will be published on the TELKOMNIKA Telecommunication Computing Electronics and Control (ISSN 1693-6930, SCOPUS/ScimagoJR indexed journal, Q2 on Electrical and Electronics Engineering, SJR: 0.283, CiteScore: 1.09, SNIP: 0.730).
Congratulations!
Please prepare your final paper (in MS Word file format) adheres every detail of the guide of authors, and check it for spelling/grammatical mistakes.

In order to cover part of the publication cost, each accepted paper will be charged:

* USD 265 (for Overseas Authors)
* IDR 3500K (for Indonesian Authors)

This charge is for the first 8 pages, and if any published manuscript over 8 pages will incur extra charges USD 50 (or IDR 600K for Indonesian Authors) per page

The payment should be made by bank transfer (T/T):

Bank Account name (please be exact)/Beneficiary: TOLE SUTIKNO
Bank Name: Bank Central Asia (BCA)
Branch Office: Kusumanegara
City: Yogyakarta
Country :Indonesia
Bank Account #: 8465122249
SWIFT Code: CENAIDJAXXX

or through PayPal (as alternative of bank transfer) to email:
tole@ee.uad.ac.id

Your paper will be scheduled for forthcoming issue. Please pay the publication fee as soon as possible (within 3 weeks). If you need more time, please send a request to this email. We can give you 5 weeks at the most.

Then, please submit your final paper & payment receipt to this email

I look forward for your response

Sincerely yours,
Dr. Tole Sutikno
Universitas Ahmad Dahlan
telkomnika@uad.ac.id

Bank's detailed address :
Bank BCA Kusumanegara
Jl. Kusumanegara No. 18
City: Yogyakarta
Province: D.I. Yogyakarta (DIY)
Country :Indonesia
Post Code: 55165
Indonesia, Phone:+62 274 418896

The Beneficiary's address:
Kampus 3 Universitas Ahmad Dahlan
Jln. Prof. Soepomo, Janturan
City: Yogyakarta
Province: D.I. Yogyakarta (DIY)
Post Code: 55164
Country: Indonesia

==========

TELKOMNIKA Telecommunication, Computing, Electronics and Control is a Scopus indexed leading journal, Scimago journal ranking (SJR) Q2 on Electrical & Electronics Engineering. This journal ONLY publishes high quality papers. A high quality paper MUST has: 1) a clear statement of the problem the paper is addressing; 2) the proposed solution(s); and 3) results achieved. It describes clearly what has been done before on the problem, and what is NEW. The goal of your final camera ready paper is to describe NOVEL TECHNICAL RESULTS.

Update title of your paper!! The title summarizes the main idea or ideas of your study (title is "summary" and "essence" of your paper). Title should be "encompassing" as well as "descriptive". A good title contains the fewest possible words needed to adequately describe the content and/or purpose of your research paper. Rarely use abbreviations or acronyms unless they are commonly known. Find the below guide, how to update your paper title.

For original research paper, there are four (4) types of novel technical results: 1) An algorithm; 2) A system construct: such as hardware design, software system, protocol, etc.; The main goal of your revised paper is to ensure that the next person who designs a system like yours doesn't make the same mistakes and takes advantage of some of your best solutions. So make sure that the hard problems (and their solutions) are discussed and the non-obvious mistakes (and how to avoid them) are discussed; 3) A performance evaluation: obtained through analyses, simulation or measurements; or 4) A theory: consisting of a collection of theorems. Your final camera ready paper should focus on: 1) Describing the results in sufficient details to establish their validity; 2) Identifying the novel aspects of the results, i.e., what new knowledge is reported and what makes it non-obvious; and 3) Identifying the significance of the results: what improvements and impact do they suggest. Number of minimum references for original research paper is 25 references (and minimum 20 recently journal articles).

For review paper, the paper should present a critical, constructive analysis of the literature in a specific field through summary, classification, analysis and comparison. The function and goal of the review paper is: 1) to organize literature; 2) to evaluate literature; 3) to identify patterns and trends in the literature; 4) to synthesize literature; or 5) to identify research gaps and recommend new research areas. The structure includes:
1. Title – in this case does not indicate that it is a review article.
2. Abstract – includes a description of subjects covered.
3. Introduction includes a description of context (paragraph 1 – 3), motivation for review (paragraph 4, sentence 1) and defines the focus (paragraph 4, sentences 2 – 3)

4. Body – structured by headings and subheadings
5. Conclusion – states the implications of the findings and an identifies possible new research fields
6. References (“Literature Review”) – organised by number in the order they were cited in the text.
Number of minimum references for review paper is 50 references (and minimum 40 recently journal articles).

Prepare your abstract in single paragraph and within 200 words. You need to summarize your contribution, idea, findings/results, and describe implications of the findings. Without abbreviations, footnotes, or references. Without mathematical equations, diagram or tabular material. It is suggested to present your abstract included the elements: 1) state the primary objective of the paper; 2) highlight the merits (or contribution; 3) give a conceptual idea on the method; 4) describe the research design and procedures/processes employed (is it simulation, experimental, survey etc.); 5) give the main outcomes or results, and the conclusions that might be drawn; and 6) include any implications for further research or application/practice, if any.
Please refer to https://bit.ly/35R6JTs and https://bit.ly/2DxU9MI for further guidelines

Your final camera ready paper should reflect a careful consideration of the following criteria:
1. Analysis: your revised paper should demonstrate a clear understanding of the key issues related to your topic of choice. The paper should display analysis and not mere summary of the topic under consideration. It should also include evidence to support arguments where necessary.
2. Connections: your paper should demonstrate a connection of the references you mention to the central topic and to each other where necessary throughout the paper.
3. Mechanics: this includes attention to punctuation, grammatical soundness and your submissions being checked for spellings errors.
4. Formatting: adhere the new guide of authors (http://icw.telkomnika.com/download/telkomnika-template-and-guide-of-authors-2020/)

Guideline for preparing your paper title:

A good research paper title: (1) Condenses the paper’s content in a few words & Use words that create a positive impression and stimulate reader interest; (2) Captures the readers’ attention; (3) Indicate accurately the subject and scope of the study and Differentiates the paper from other papers of the same subject area. Five (5) Simple steps to write a good research paper title:
STEP 1: Ask yourself these questions and make note of the answers:
- What is my paper about?
- What techniques/ designs were used?
- Who/what is studied?
- What were the results?

STEP 2: Use your answers to list key words.
STEP 3: Create a sentence that includes the key words you listed
STEP 4: Delete all unnecessary/ repetitive words and link the remaining.
STEP 5: Delete non-essential information and reword the title.

General Guidelines:
---
1. Please re-read our instructions (at: http://journal.uad.ac.id/index.php/TELKOMNIKA/about/editorialPolicies#custom-1) carefully and follow the checklist strictly, as any spelling mistakes and errors may be translated into the typeset version.
2. The “result and discussion” section reports the most important findings, including analysing results as appropriate. It is very important to prove that your manuscript has a significant value and not trivial.
3. Please re-check that all references are already cited in your article, and order of your citation is SEQUENTIAL

example in a paper:
—> [1-4], [2], [5-6], [7-9], [8], [4-5], [9], [10-14], [12], [15] .....
(SEQUENTIAL) — correct
—> [1], [2], [3], [4-6], [7], [8-10], [4-5], [11-16], [13], [17] .....
(SEQUENTIAL) — correct
—> [2], [3], [6], [1], [4-5], [7], [11-16], [8-10], [13], [17] ..... (NOT SEQUENTIAL) — INCORRECT
—> [4-5], [7], [2], [3], [17], [6], [1], [11-16], [8-10], [13] ..... (NOT SEQUENTIAL) — INCORRECT

4. If you need references to enrich your references and improve your paper, please get take a look at:
- http://beei.org
- http://ijece.iaescore.com
- http://ijeecs.iaescore.com
- http://journal.uad.ac.id/index.php/telkomnika
- http://ijpeds.iaescore.com
- http://ijai.iaescore.com

Please use “Search” at menu "Journal Content" in right side of the site.

TELKOMNIKA (Telecommunication Computing Electronics and Control)
http://www.journal.uad.ac.id/index.php/TELKOMNIKA

Hartono Ibbi <hartonoibbi@gmail.com> 25 Desember 2019 pukul 13.58
Kepada: "Dr. Tole Sutikno" <telkomnika@uad.ac.id>

Dear Dr. Tole Sutikno
Editor in Chief of TELKOMNIKA

Thank you for the opportunity given to us to publish our research results at TELKOMNIKA. We will make the payment process immediately and likewise we will also improve our writing based on input from editors, reviewers, and mentors given at the conference.

Happy to join ICW-TELKMONIKA 2019 and hope to participate in CITEI 2020 and ICW-TELKMONIKA 2020. Thank you again for the good service during the ICW-TELKOMNIKA 2019 and if there is anything else we need to complete, don't hesitate to let us know.

Best Regards,

Hartono
[Kutipan teks disembunyikan]

Jurnal TELKOMNIKA <telkomnika@uad.ac.id> 25 Desember 2019 pukul 13.58
Kepada: hartonoibbi@gmail.com

Dear Sir/Madam,

Your email has reached my queues safely, so i will get back in touch with you there as soon as i can. Just a heads up i might not get back to you straight away, as i answer all emails in order to keep things fair.

How to submit your manuscript
To make a submission, you must have a user account and be enrolled as an Author. User accounts can either be created by the Journal Manager or you can register yourself (this journal policy allow you create user account by your self as a Reader, an Author and/or a Reviewer). All fields with an
asterisk beside them (Username; Password; Repeat Password; First Name; Last Name; Email) are mandatory. Your username and your email address must be unique; furthermore, while you can change your email address at a later date, you will be unable to change your username. If you want to register
in another role within the same journal (for example, if you are already a Reader, but also want to become an Author) you can log in; go to Edit My Profile (under My Account on your User Home page); and check off the check-boxes next to any available roles, near the bottom of the page. Once you have an account, log in to the journal site and select the role of Author. The Author is asked to upload a submission file and to provide metadata or indexing information. (The metadata improves the search
capacity for research online and for the journal.) The Author can upload Supplementary Files, in the form of data sets, research instruments, or source texts that will enrich the item, as well as contribute to more open and robust forms of research and scholarship.

Your paper ID number?
Please ALWAYS put your paper ID number in the subject line of email communication when making inquiries. For example, if your submitted manuscript URL:
http://journal.uad.ac.id/index.php/TELKOMNIKA/author/submission/1234, your paper ID number is 1234.

Your cooperation is very appreciated.

Thank you

Best Regards,
Tole Sutikno
Editor:
TELKOMNIKA TCEC, http://journal.uad.ac.id/index.php/TELKOMNIKA

Journal Form for Hardcopy Order

Please download Journal form for Hardcopy order http://goo.gl/a6Y7zg and contact us if further information is needed: Mr. Sulchan Hidayat: sulchan.hidayat@gmail.com, +62 81328731075

--

Best Regards,
Tole Sutikno
Editor-in-Chief, TELKOMNIKA (Telecommunication, Computing, Electronics and Control)
ISSN: 1693-6930, e-ISSN: 2302-9293
http://journal.uad.ac.id/index.php/TELKOMNIKA

UNIVERSITAS AHMAD DAHLAN

Kampus 1: *Jln. Kapas No. 9 Yogyakarta*

Kampus 2: *Jl. Pramuka 42, Sidikan, Umbulharjo, Yogyakarta 55161*

Kampus 3: *Jl. Prof. Dr. Soepomo, S.H., Janturan, Warungboto, Umbulharjo, Yogyakarta 55164*

Kampus 4: *Jl.Ringroad Selatan, Yogyakarta*

Kampus 5: *Jl. Ki Ageng Pemanahan 19, Yogyakarta*

Kontak

Email: info@uad.ac.id

Telp. : (0274) 563515, 511830, 379418, 371120

Fax. : (0274) 564604

Hartono Ibbi <hartonoibbi@gmail.com> 7 Januari 2020 pukul 13.45
Kepada: "Dr. Tole Sutikno" <telkomnika@uad.ac.id>

Dear Dr. Tole Sutikno
Editor In Chief of Telkomnika

We have pay the publication fee of Rp. 3,500,000, which can be seen in this email attachment. Then we also want to inform you that we are in the process of revising our paper and it is hoped that this will be completed before January 20, 2020. Thank you for the opportunity given to us to publish our research on TELKOMNIKA.

Best Regards,

Hartono

Pada tanggal Min, 22 Des 2019 pukul 21.26 Tole Sutikno <telkomnika@uad.ac.id> menulis:
[Kutipan teks disembunyikan]

Bukti Bayar Telkomnika.jpg
1371K

Jurnal TELKOMNIKA <telkomnika@uad.ac.id> 7 Januari 2020 pukul 13.45
Kepada: hartonoibbi@gmail.com

[Kutipan teks disembunyikan]

Revisi Sesuai Masukan Reviewer

Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) Method for multi-class imbalanced datasets

H Hartono*[1,2], Erianto Ongko[3], Dahlan Abdullah[4], Yeni Risyani[5]
[1]Department of Computer Science, STMIK IBBI, Medan, Indonesia
[2]Department of Computer Science Universitas Potensi Utama, Medan, Indonesia
[3]Department of Informatics, Akademi Teknologi Industri Immanuel, Medan, Indonesia
[4]Department of Informatics, Universitas Malikussaleh, Medan, Indonesia
[5]Department of Information Systems, STMIK IBBI, Medan, Indonesia
*Corresponding author, e-mail: hartonoibbi@gmail.com[1,2], eriantoongko@gmail.com[3], dahlan@unimal.ac.id[4], ms_yenir@yahoo.com[5]

Abstract

Research on multi-class imbalance from a number of researchers faces obstacles in the form of poor data diversity and a large number of classifiers. The Hybrid Approach Redefinition-Multiclass Imbalance (HAR-MI) method is a Hybrid Ensembles method which is the development of the Hybrid Approach Redefinion (HAR) method. This study has compared the results obtained with the Dynamice Ensemble Selection-Multiclass Imbalance (DES-MI) method in handling multiclass imbalance. In the HAR-MI Method, the preprocessing stage was carried out using the Random Balance Ensembles Method and Dynamic Ensemble Selection to produce a candidate ensemble and the processing stages was carried out using Different Contribution Sampling and Dynamic Ensemble Selection to produce a candidate ensemble. This research has been conducted by using multi-class imbalance datasets sourced from the KEEL Repository. The results show that the HAR-MI Method can overcome good multi-class imbalance with better data diversity, smaller number of classifiers, and better classifier performance compared to a DES-MI Method.

Keywords*: multi-class imbalance, Data Diversity, Classifier, Hybrid Approach Redefinition-Multiclass Imbalance*

Copyright © 2020 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Class imbalance occurs if a class or several classes become underrepresented so it is also called a minority class because it has instances that are much smaller than other classes[1]. In machine learning research, class imbalance problems are the main challenges that attract the attention of a number of researchers[2]. Research on this issue is included in the 20 main research topics that are the most interesting in machine learning, especially big data. Minority Class is also called a positive class because it is a class with interesting patterns to observe. For comparison, the detection of breast cancer sufferers is often a class with a small number of instances, if the classification process for detection of breast cancer sufferers experiences class imbalance problems then there is a possibility that detection of patients is not obtained even though the sufferer class is very interesting to obtain[3].

There are a number of methods that have been proposed to deal with class imbalance problems such as resampling, cost sensitive, ensemble learning, kernel-based methods, and active learning methods[4]. Multi-class imbalance problems are far more complicated to handle than two-class imbalances. The multi-class imbalance condition will be more difficult if the desired results are as accurate as possible in accordance with the existing problem. On the other hand, applying the method proposed to handle two-class imbalance problems to handle multi-class imbalance problems does not get the desired results [5]. In general, the algorithm for handling multi-class imbalance is to develop an algorithm used for handling binary class Imbalace through the decomposition method[6]. Another common method is to adopt an ensemble-based approach for use in handling multi-class imbalances[4] and another way is to adapt the intrigue process by building decision trees[7]. A relatively easy way to do is to view multi-class imbalance as a subset of binary problems[8][9].

The multi-class imbalance problems that will be solved are problems such as Many Minority-One Majority, One Minority-Many Majority, and Many Minoriy-Many Majority[10]. Zhang, Luo, García, Tang, & Herrera [1] suggested that to overcome the problem of imbalance class

there are 2 (two) things that need to be considered, namely those related to the number of classifiers and diversity (diversity) of data. [11] propose the Dynamic Classifier Selection (DCS) method for dealing with multi-class imbalance problems, but it has the disadvantage of being a large number of classifiers. García, Zhang, Altalhi, Alshomrani, & Herrera suggested the Dynamic Ensemble Selection (DES) -MI method which gives better results compared to the Dynamic Classifier Selection (DCS) method. The DES-MI method found has a small classifier, but in research conducted by Pérez-Gállego, Castaño, Ramón Quevedo, & José del Coz [12]has identified that diversity data obtained by DES-MI is not good enough. The Hybrid Approach Redefinition (HAR) method which is a Hybrid Esembles approach can overcome the problem of cass imbalance with a small number of classifiers and good data diversity, on two-class imbalance problems[13][14].

This research will optimize the HAR method so that it can be used to overcome multi-class imbalance problems. In the optimization process the preprocessing stages were carried out using the Random Balance Ensemble Method proposed by Díez-Pastor, Rodríguez, García-Osorio, & Kuncheva [15] and Dynamic Ensemble Selection so that a candidate ensemble on multiclass problems and processing stages was carried out using Different Contribution Sampling proposed by Jian, Gao, & Ao [16] and Dynamic Ensemble Selection. This research will be conducted using multi-class imbalanced datasets sourced from the KEEL Repository [17]. The results of the study are expected to be able to overcome the problem of multi-class imbalance with better data diversity and the small number of classifiers.

Comment [Office1]: Analysis: your revised paper should demonstrate a clear understanding of the key issues related to your topic of choice. The paper should display analysis and not mere summary of the topic under considerati

Comment [Office2]: Connections: your paper should demonstrate a connection of the references you mention to the central topic and to each other where necessary throughout the paper.

2. Research Method

This research will produce the HAR-MI method to overcome multi-class imbalance problems. HAR Method will be carried out an optimization process with HAR-MI Method so that it can handle multi-class imbalance problems by adding capabilities from HAR Method to determine candidate ensembles by using Dynamic Ensemble Selection on minority classes and majority classes so that they can recognize each subset of minority and majority classes based on 2-Dimensional Datasets proposed by Sáez et al [10]. The results of HAR-MI Method are expected to obtain better data diversity and also a small number of classifers.

The stages of research conducted by researchers from this study can be seen in Figure 1.

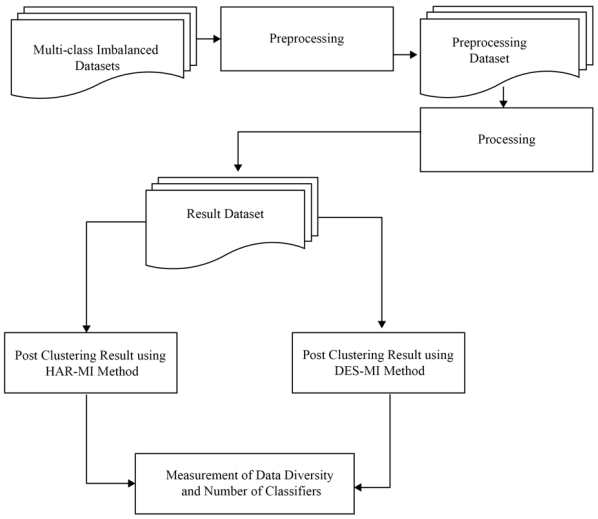

Figure 1. Stages of Research Methods

In Figure 1, it can be seen that the process that occurs in the dataset selection and preparation stage is determined by the imbalance dataset with varying imbalance ratio. The next process is Preprocessing. The process of handling the multi-class imbalance will begin with the preprocessing stage. The purpose of this preprocessing stage is to reduce the number of classifiers. Where the preprocessing stage will be done using the Random Balance Ensemble Method method and Dynamic Ensemble Selection. The Random Balance Ensemble Method will use Random Under Sampling and SMOTEBoost. The results of the preprocessing stage are in the form of a preprocessing dataset which will then proceed to the processing stage.

Preprocessing Datasets that still have imbalance class problems will enter the Processing stage. this processing stage is done by using the Different Contribution Sampling (DCS) method and Dynamic Ensemble Selection.

2.1. Multi-Class Imbalance Problem

Multi-class imbalance issues are more complicated to handle than two-class imbalances, because decision boundaries involve separation for more classes. Applying the method proposed to deal with two-class imbalance problems to deal with multi-class imbalance problems will not get the desired results [18]. The multi-class imbalance problem that will be solved is the problem involving: Many Minority-One Majority, One Minority-Many Majority, and Many Minori-Many Majority [10].

Napierala & Stefanowski [19] suggested a framework that can be used to facilitate understanding of multi-class imbalance in the form of 2-dimensional datasets as can be seen in Figure 2 [10].

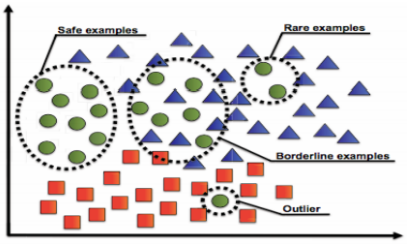

Figure 2. *2-Dimensional Dataset*

In Figure 2.4 it can be seen that the Majority Class in general can be divided into 2 (two) symbols, namely: Δ and □. Where □ has a value of cn (e) or neighborhood value that is greater than Δ. While the Minority Class uses symbols O and is divided into: Safe Examples, Borderline Examples, Rare Examples, and Outliers. Safe Examples have a greater cn (e) value compared to Borderline Examples, Borderline Examples have a greater cn (e) value compared to Rare Examples, and Rare Examples have a value of cn (e) that is greater than outliers.

2.2. Preprocessing and Processing Stage in HAR-MI method

The preprocessing stage was carried out using the Random Balance Ensembles Method and Dynamic Ensemble Selection. The pseudocode of this stage is as follows.

Require: Set S of examples(x_1,y_1)
Ensure: New set *S'* of examples with *Random Balance* and Dynamice Ensemble Selection
1: *totalSize*←|S|
2: Determine *k* as the number of *Nearest Neighbor*
3: For All Samples in S do
4: Determine the Borderline of Positive or Minority Class as $E_O C_t^+$
5: Determine the Borderline of Negative or Majority Class as $E_O C_t^-$
6: End For
7: For All Samples in $E_O C_t^+$ do
8: Calculate the $cn(e)_i$ as neigborhood value for each sample
9: Order Ascending the sample according to the $cn(e)_i$
10: End For

11: Building a candidate ensemble for *Safe*, *Borderline*, *Rare*, dan *Outlier* according to *k* value
12: Take a candidate ensemble of *Safe*, *Borderline*, *Rare*, dan *Outlier* to *SP*
13: For All Samples in $E_O C_t^-$ do
14: Take a candidate ensemble to SN
15: End For
16: Add Instance from with S|yi=+1 to S_P
17: Add Instance from with S|yi=-1 to S_N
18: Calculate the size of Majority Class from S_N
19: Calculate the size of Majority Class from S_P
20: *newMajoritySize*←Random integer between 2 and *totalSize*-2
21: *newMinoritySize*←*totalSize* – *newMajoritySize* 8: if newMajoritySize
22: **if** *newMajoritySize* <*majoritySize* **then**
23: $S^{'} \leftarrow S_P$
24: $S^{'}$ will fill with a random instance from S_N
25: Create *newMinoritySize* – *minoritySize* artificial
26: **else**
27: $S^{'} \leftarrow S_N$
28: $S^{'}$ will fill with a random instance from S_P
29: create *newMajoritySize* – *majoritySize* artificial
30: **end if**
31: **return** $S^{'}$

Based on the pseudocode above, it can be seen that in the preprocessing stage Random Balance Ensemble Method was carried out using Random Under Sampling and SMOTEBoost. In the Random Undersampling process the Dynamic Ensemble Selection process will take the form of borderline determination for minority and majority class. Then for samples that are in the borderline minority class $E_O C_t^+$, the neighborhood value calculation process cn (e) will be performed, then it will be sorted ascending to determine the candidate ensemble for Safe, Borderline, Rare, and Outlier, then the candidate ensemble will be included in the *SP*. Next for the sample that is in the borderline the major class will be entered into *SN*.

After that, the process will continue with the Random Balance Ensemble Method, which will be based on the results of the Dynamic Ensemble Selection. The process starts with the determination of Majority and Minority Size. Then based on the determination of the size, an imbalance class will be handled. If the size of the new Majority Class is greater than the new Majority Class, this means that the Minority Class is larger than the Majority Class and part of the Minority Class instance will be taken to move to the Majority Class and vice versa. Determination of the sample will be done by Random Under Sampling and the determination of the instance that will be transferred will be done with SMOTEBoost.

The Processing stages was carried out using the Different Contribution Sampling and Dynamic Ensemble Selection. The pseudocode of this stage is as follows.

1: **Input**: *S: Training Set; T:* Number of Iterations*; n*: Bootstrap Size*; k*: *neighbors*
2: **Output:** Bagged Classifier: $H(x) = sign\left(\sum_{t=1}^{T} h_t(x)\right)$ where h_t[-1, 1] are the induced classifiers
3: **Process:**
4: For All Samples in S do
5: Determine the Borderline of Positive or Minority Class as $E_O C_t^+$
6: Determine the Borderline of Negative or Majority Class as $E_O C_t^-$
7: End For
8: For All Samples in $E_O C_t^+$ do
9: Calculate the $cn(e)_i$ as neigborhood value for each sample
10: Order Ascending the sample according to the $cn(e)_i$
11: End For
12: Building a candidate ensemble for *Safe*, *Borderline*, *Rare*, dan *Outlier* according to *k* value
13: Take a candidate ensemble of *Safe*, *Borderline*, *Rare*, dan *Outlier* to *SP*
14: For All Samples in $E_O C_t^-$ do
15: Take a candidate ensemble to SN
16: End For
17: **for** *i* = 1 to Number of Instance in Preprocessed Dataset **do**
18: Add Preprocessed Dataset to S_i

19 B-SVM will do for classifying S_i
20: Determine the Majority Class
21: Determine the Minority Class
22: For All Instance in Majority Class do
23: NewSVSets[] will form by checking and delete the noise in SV Sets
24: NewNSVSets[]will form by multiple RUS
25: end while
26: For All instance from new SV Sets and NSV do
27: Create an instance for Majority Class
28: End For
29: For All Instance in Minority Class do
30: SMOTEBoost Process for SV Sets and create SMOTESets
31: end while
32: For All SMOTESets and NewNSVSets do
33: New PositiveSampleSets
34: End For
35: For All NewNegativeSampleSets and NewPositiveSampleSets do
36: ResultDataSet
37: End For
38: End For

After the preprocessing dataset is generated, the Dynamic Ensemble Selection process will occur at the initial stage for borderline determination of minority and majority class. Then the next step will be the Differential Contribution Sampling process where both majority classes and minority classes will be divided into SV Sets and NSV Sets. NSV Sets in the Negative Sample will undergo a Multiple RUS process, while SV Sets in the Positive Sample will experience a SMOTEBoost.

2.3. Data Diversity

In the ensemble learning process, in reality if there is a classifier that can guarantee that there is no misclassification, an ensemble process is not needed on the classifier. The ensemble process in the classifier occurs in the hope that better results can be obtained. Assuming that if there is a misclassification of the classifier in a part it can be covered by merging with other classifiers that also missclassification in other parts[20].

According to Díez-Pastor, Rodríguez, García-Osorio, & Kuncheva [15] it is important to pay attention to the diversity of data in handling imbalance classes. This means that attempted misclassification produced by each classifier is as small as possible and if there is misclassification it is expected to occur on different objects or parts [20].

Suppose that Z= $\{z_1, \ldots, z_n\}$ which is a dataset that is in the decision region $\Re^n$, so that $z_j \in \Re^n$ it is an instance involved in the classification problem. Then the output of the classifier D_i as a classifier paired comparison matrix (relationship pairwise classifier) can be seen in Table 1.

Table 1. Relationship Pairwise Classifier Matrix[20]

	D_k *Correct* (1)	D_k *Wrong* (0)
D_i *Correct* (1)	N^{11}	N^{10}
D_i *Wrong* (0)	N^{01}	N^{10}

Diversity data can be calculated using Q-Statistics[21].

$$Q_{i,k} = \frac{N^{11}N^{00} - N^{01}N^{10}}{N^{11}N^{00} + N^{01}N^{10}} \quad (1)$$

2.4. Classifier

Classifiers can generally be defined as *Decision Region* $\Re^n$ that place an object into a set *class* Ω, where Ω consists of class ω_1, ω_2, until ω_n. This can be seen in Equation 9[20].

$$D: \Re^n \rightarrow \Omega \quad (2)$$

Where D is the classifier and is the set of each point in the decision region $\Re^i$ which is intended for *class* ω_i.

2.5. Classifier Performance

ROC Curve is one statistical method that is often used to determine the performance of a classifier. This curve is generated by plotting the true positive fraction of a positive sample in the Y axis with the false positive fraction of a negative sample (False Positive Rate) in the X axis[22]. The concepts of True Positive and False Positive can be seen in the Confusion Matrix as can be seen in Table 2.[23].

Table 2. Confusion Matrix[24]

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

The number of performance classifier measurement parameters in the two class problems are as follows[25].

$$T\,Prate = \frac{\mathrm{TP}}{\mathrm{TP + FN}} \quad (3)$$

$$F\,Prate = \frac{\mathrm{FP}}{\mathrm{TN+FP}} \quad (4)$$

$$TN\ rate\ = \frac{\mathrm{TN}}{\mathrm{TN + FP}} \quad (5)$$

$$\underline{R}ecall = TPrate \quad (6)$$

$$\underline{P}recision = PPValue = \frac{\mathrm{TP}}{\mathrm{TP + FP}} \quad (7)$$

$$F\text{-}Measure = \frac{2\mathrm{RP}}{\mathrm{R+P}} \quad (8)$$

$$G\text{-}Mean = \sqrt{TPrate\ .\ TNrate} \quad (9)$$

True Positive Rate (TPrate) is stated as a recall which states the percentage of data captured is relevant data. Positive Predictive Value (PPValue) is stated as Precision which states the percentage of relevant data identified to be taken. F-Measure states the harmonic average value between recall and precision. The F-Measure value is usually smaller than 2, the higher the value of F-Measure states that both recall and precision are quite high. G-Means on the other hand states the balance between positive samples (minority class) and negative samples (majority class)[23].

Performance measurement in multi class imbalance is basically a modification of two class problems, and in general there are 2 (two) parameters used, namely:*MAvA* and *MFM*[26].

$$MAvA = \frac{\sum_{i=1}^{N} ACC_i}{m} \quad (10)$$

where *m* is the number of classes and ACC_i stands for the accuracy rate for the class I and *MAvA* is the average value of accuracy.

$$MFM = \frac{F-measure_i}{m} \quad (11)$$

where *MFM* is the muti class F-Measure

3. Results and Analysis

3.1. Dataset Description

This study uses a multi-class imbalanced dataset that is sourced from the KEEL Repository. The dataset selected in this study has represented a low, medium and high imbalance ratio. For datasets with a low imbalance ratio are Balance Scale datasets, datasets with moderate imbalance ratio are Car Evaluaton datasets, and dataset with high imbalance ratio are Red Wine Quality datasets, Ecoli, and Pageblocks. Dataset description can be seen in Tabe 3[17].

Table 3. Dataset Description[17]

Dataset	**#Ex**	**#Atts**	**Distribution of Class**	**IR**
Balance Scale	625	4	288/49/288	5.88
Car Evaluation	1728	6	384/69/1210/65	18.62
Red Wine Quality	1599	11	10/53/681/638/199/18	68.1
Ecoli	336	7	2/2/5/20/35/52/77/143	71.5
Pageblocks	548	10	3/8/12/33/492	164

3.2. Testing Result

The first test is to obtain a comparison of the number of classifier and diversity data obtained by using HAR-MI and DES-MI Method. Testing of each method will be carried out as many as 10 testing for each dataset. The average test results can be seen in Table 4.

Table 4. Testing Result for Number of Classifier and Data Diversity for Each Method

Dataset	HAR-MI Method		DES-MI Method	
	Number of Classifier	*Data Diversity (Q-Statistics)*	*Number of Classifier*	*Data Diversity* (Q-Statistics)
Balance Scale	191.6	0.397	197.2	0.421
Car Evaluation	471.6	0.457	487.9	0.461
Red Wine Quality	397.8	0.431	395.3	0.411
Ecoli	91.1	0.397	121.2	0.413
Pageblocks	117.8	0.441	119.6	0.456

Based on the results in Table 4, it can be seen that HAR-MI Method gives better results on better data diversity in the three datasets when compared with DES-MI Method. The test results for the HAR-MI Method classifier are better in the Balance Scale, Car Evaluation, Ecoli, and Pageblocks datasets. For the Red Wine Quality dataset, DES-MI is slightly superior compared to HAR-MI. There is a tendency if the number of attributes increases, the sampling process, especially on Random Under Sampling, requires a larger classifier. However, the difference in the number of classifiers is not very significant.

The results of testing *MAvA* and *MFM* can be seen in Table 5.

Table 5. Testing Result for *MAvA* and *MFM* for Each Method

Dataset	HAR-MI Method		DES-MI Method	
	MAvA	*MFM*	*MAvA*	*MFM*
Balance Scale	66.71	0.71	61.29	0.61
Car Evaluation	97.68	0.97	94.27	0.945
Red Wine Quality	45.24	0.43	41.81	0.395
Ecoli	57.31	0.58	49.67	0.51
Pageblocks	47.81	0.49	45.92	0.44

In Table 5 it can be seen that HAR-MI Method gives better results for MAVA and MFM when compared to DES-MI Method. Both methods have provided excellent MAVA and MFM values. A good MAVA means that the accuracy of the classification has been very good, where the misclassification that occurs is very minimal. This means that the instance of the minority class has been classified correctly and also the majority class instances that are incorrectly classified as minority classes are also minimal. This is because F-Measure states how many instances in the minority class are correctly defined and also measures how many instances in the majority class are incorrectly classified as minority classes.

4. Conclusion

Based on the test results it can be seen that HAR-MI Method gives better results compared to DES-MI Method for both the number of classifier, data diversity, and also the performance classifier. It should be noted that for the number of classifiers, where if the dataset has many attributes such as the Red Wine Quality, then the HAR-MI Method can produce poor results. In general, the imbalance ratio does not have a significant effect on the test results. This means that both HAR-MI Method and DES-MI Method can handle the imbalance problem class very well. Future research, it is expected that HAR-MI Method can be optimized so that it can be applied to datasets for a large number of attributes without causing a large number of classifiers. The main attention needs to be given to the sampling method used in the HAR-MI Method. It is necessary to find another sampling alternative at the preprocessing and processing stages.

Acknowledgment

This work was supported by the Grant of Ministry of Research, Technology, and Higher Education (KEMENRISTEKDIKTI) of the Republic of Indonesia.

References

[1] Shuo Wang, Xin Yao. Multiclass Imbalance Problems: Analysis and Potential Solutions. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*. 2012 Aug;42(4):1119–30.

[2] Krawczyk B. Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence*. 2016 Nov;5(4):221–32.
[3] Ali A, Shamsuddin SM, Ralescu AL. Classification with class imbalance problem: A Review. *International Journal of Advances in Soft Computing and Its Application*. 2015;7(3):176–204.
[4] Haixiang G, Yijing L, Shang J, Mingyun G, Yuanyue H, Bing G. Learning from Class-Imbalanced Data: Review of Methods and Applications. *Expert Systems With Applications*. 2017 May 1;73:220–39.
[5] Fernández A, López V, Galar M, del Jesus MJ, Herrera F. Analysing the classification of imbalanced data-sets with multiple classes: Binarization techniques and ad-hoc approaches. *Knowledge-Based Systems*. 2013 Apr 1;42:97–110.
[6] Bi J, Zhang C. An empirical comparison on state-of-the-art multi-class imbalance learning algorithms and a new diversified ensemble learning scheme. *Knowledge-Based Systems*. 2018 Oct 15;158:81–93.
[7] Hoens TR, Qian Q, Chawla NV, Zhou Z-H. Building Decision Trees for the Multi-class Imbalance Problem. In: Tan P-N, Chawla S, Ho CK, Bailey J, editors. *Advances in Knowledge Discovery and Data Mining*. Springer Berlin Heidelberg; 2012. p. 122–34. (Lecture Notes in Computer Science).
[8] Allwein EL, Schapire RE, Singer Y. Reducing Multiclass to Binary: A Unifying Approach for Margin Classifiers. *Journal of Machine Learning Research*. 2000;1(Dec):113–41.
[9] Galar M, Fernández A, Barrenechea E, Bustince H, Herrera F. An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes. *Pattern Recognition*. 2011 Aug 1;44(8):1761–76.
[10] Sáez JA, Krawczyk B, Woźniak M. Analyzing the oversampling of different classes and types of examples in multi-class imbalanced datasets. *Pattern Recognition*. 2016 Sep;57:164–78.
[11] Zhang Z-L, Luo X-G, García S, Tang J-F, Herrera F. Exploring the effectiveness of dynamic ensemble selection in the one-versus-one scheme. *Knowledge-Based Systems*. 2017 Jun 1;125:53–63.
[12] Pérez-Gállego P, Castaño A, Ramón Quevedo J, José del Coz J. Dynamic ensemble selection for quantification tasks. *Information Fusion*. 2019 Jan 1;45:1–15.
[13] Hartono, Ongko E, Sitompul OS, Tulus, Nababan EB, Abdullah D. Hybrid Approach Redefinition (HAR) Method with Loss Factors in Handling Class Imbalance Problem. In: 2018 International Symposium on Advanced Intelligent Informatics (SAIN). 2018. p. 56–61.
[14] Hartono, Sitompul OS, Nababan EB, Tulus, Abdullah D, Ahmar AS. A New Diversity Technique for Imbalance Learning Ensembles. *International Journal of Engineering & Technology*. 2018 Apr 8;7(2):478–83.
[15] Díez-Pastor JF, Rodríguez JJ, García-Osorio CI, Kuncheva LI. Diversity techniques improve the performance of the best imbalance learning ensembles. *Information Sciences*. 2015 Dec;325:98–117.
[16] Jian C, Gao J, Ao Y. A new sampling method for classifying imbalanced data based on support vector machine ensemble. *Neurocomputing*. 2016 Jun;193:115–22.
[17] Alcalá-Fdez J, Sánchez L, García S, Jesus MJ del, Ventura S, Garrell JM, et al. KEEL: a software tool to assess evolutionary algorithms for data mining problems. *Soft Comput*. 2009 Feb 1;13(3):307–18.
[18] Fernandez E, Navarro J, Duarte A, Ibarra G. Core: A decision support system for regional competitiveness analysis based on multi-criteria sorting. *Decision Support Systems*. 2013 Feb 1;54(3):1417–26.
[19] Napierala K, Stefanowski J. Identification of Different Types of Minority Class Examples in Imbalanced Data. In: Hybrid Artificial Intelligent Systems [Internet]. Springer, Berlin, Heidelberg; 2012 [cited 2018 Aug 6]. p. 139–50. (Lecture Notes in Computer Science). Available from: https://link.springer.com/chapter/10.1007/978-3-642-28931-6_14
[20] Kuncheva LI. Combining Pattern Classifiers. John Wiley & Sons; 2004. 295–327 p.
[21] Kuncheva LI, Whitaker CJ. Measures of Diversity in Classifier Ensembles and Their Relationship with the Ensemble Accuracy. *Machine Learning*. 2003 May 1;51(2):181–207.
[22] Gigliarano C, Figini S, Muliere P. Making classifier performance comparisons when ROC curves intersect. *Computational Statistics & Data Analysis*. 2014 Sep 1;77:300–12.
[23] Sun Y, Kamel MS, Wong AKC, Wang Y. Cost-sensitive boosting for classification of imbalanced data. *Pattern Recognition*. 2007 Dec;40(12):3358–78.
[24] Zhang L, Yang H, Jiang Z. Imbalanced biomedical data classification using self-adaptive multilayer ELM combined with dynamic GAN. *Biomed Eng Online* [Internet]. 2018 Dec 4 [cited 2019 Oct 19];17. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6280414/
[25] Tharwat A. Classification assessment methods. *Applied Computing and Informatics* [Internet]. 2018 Aug 21 [cited 2019 Oct 19]; Available from: http://www.sciencedirect.com/science/article/pii/S2210832718301546
[26] Alejo R, Antonio JA, Valdovinos RM, Pacheco-Sánchez JH. Assessments Metrics for Multi-class Imbalance Learning: A Preliminary Study. In: Carrasco-Ochoa JA, Martínez-Trinidad JF, Rodríguez JS, di Baja GS, editors. Pattern Recognition. Springer Berlin Heidelberg; 2013. p. 335–43. (Lecture Notes in Computer Science).

Proof Reading

[TELKOMNIKA Proofreading] APRIL 2020 Paper ID 14818

5 pesan

Hany Safitry F <hanyjournal.iaes@gmail.com> 19 Februari 2020 pukul 15.10
Kepada: hartonoibbi@gmail.com, eriantoongko@gmail.com, dahlan@unimal.ac.id, ms_yenir@yahoo.com

Dear Sir/Madam,

We hope this email reaches you well. We are from TELKOMNIKA Staff want to inform you that your paper **ID 14818** entitled **HAR-MI method for multi-class imbalanced datasets authentication** will be published in **April 2020 Issue**. We suggest you to checking your paper that attaches in this email.

If you have something to tell us about your paper, please reply to this email and comment on your pdf. We wait for your responses. We will wait for your responses until **February 20, 2020**. If you have passed the deadline, then we consider it to be the final paper. Thank you for your cooperation.

Best Regards,
Hany Safitry.
TELKOMNIKA Staff

TELKOMNIKA Journal
Telecommunication, Computing, Electronics, and Control

Editor in Chief : telkomnika@uad.ac.id
Staff : team@journal.uad.ac.id cc: hanyjournal.iaes@gmail.com/mitadewi.iaes@gmail.com

Universitas Ahmad Dahlan, 4th Campus, 9th Floor, LPPI Room
Ringroad Selatan St., Kragilan, Tamanan, Banguntapan, Bantul, Yogyakarta, Indonesia 55191

Website : journal.uad.ac.id/index.php/TELKOMNIKA/index

29 14818 (11jul19) (19jan20) (23feb20) Ezha.pdf
330K

Hartono Ibbi <hartonoibbi@gmail.com> 22 Februari 2020 pukul 13.50
Kepada: Hany Safitry F <hanyjournal.iaes@gmail.com>

Dear Hany Safitri,

Sorry, due to my busy schedule taking care of my parents who are sick, I just had time to reply to your email. I have received your email well and I feel the contents of the article are appropriate and can be published. Once again, I apologize for the delay in responding to your email because of my busy caring for parents.

Best Regards,

Hartono
[Kutipan teks disembunyikan]

Hany Safitry F <hanyjournal.iaes@gmail.com> 23 Februari 2020 pukul 14.56
Kepada: Hartono Ibbi <hartonoibbi@gmail.com>

Dear Sir/Madam,

Thank you for your confirmation. We will upload the pdf to the system.
[Kutipan teks disembunyikan]
[Kutipan teks disembunyikan]

[Kutipan teks disembunyikan]

Hartono Ibbi <hartonoibbi@gmail.com> 24 Februari 2020 pukul 08.00
Kepada: Hany Safitry F <hanyjournal.iaes@gmail.com>

Dear Hany Safitri,

Thank you for your help. Please do not hesitate to let us know, if there is anything else we need to complete. Thank You.

Best Regards,

Hartono
[Kutipan teks disembunyikan]

Hany Safitry F <hanyjournal.iaes@gmail.com> 24 Februari 2020 pukul 10.48
Kepada: Hartono Ibbi <hartonoibbi@gmail.com>

Dear Sir/Madam,

Thank you for your confirmation. We will upload the pdf on the system
[Kutipan teks disembunyikan]
[Kutipan teks disembunyikan]

[Kutipan teks disembunyikan]