18-Feb-2023  
  
Dear Mr. Nguyen:  
  
I have now reviewed the recommendation of the associate editor and reviewer comments on your manuscript TAI-2023-Jan-A-00016 entitled "SIMPOR: Synthetic Information towards Maximum Posterior Ratio for deep learning on Imbalanced Data" which you submitted to the IEEE Transactions on Artificial Intelligence.  
  
On the basis of this information, your manuscript cannot be accepted in its current form for publication in the IEEE Transactions on Artificial Intelligence.  
  
The associate editor and reviewer comments identified the key issues that require your attention when revising the manuscript. These comments are included at the bottom of this letter.  
  
A thorough revision of your work could result in a manuscript meriting further evaluation by the Journal. Thus I encourage you to submit a revision.  
  
To revise your manuscript, log into <https://mc.manuscriptcentral.com/tai-ieee> and enter your Author Center, where you will find your manuscript title listed under "Manuscripts with Decisions."  Under "Actions," click on "Create a Revision."  Your manuscript number has been appended to denote a revision.  
  
You will be unable to make your revisions on the originally submitted version of the manuscript.  Instead, revise your manuscript using a word processing program and save it on your computer.  Please also highlight the changes to your manuscript within the document by using the track changes mode in MS Word or by using bold or colored text.  
  
Once the revised manuscript is prepared, you can upload it and submit it through your Author Center.  
  
When submitting your revised manuscript, you will be able to respond to the comments made by the reviewer(s) in the space provided.  You can use this space to document any changes you make to the original manuscript.  In order to expedite the processing of the revised manuscript, please be as specific as possible in your response to the reviewer(s).  
  
IMPORTANT:  Your original files are available to you when you upload your revised manuscript.  Please delete any redundant files before completing the submission.  
  
Because we are trying to facilitate timely publication of manuscripts submitted to the IEEE Transactions on Artificial Intelligence, your revised manuscript should be submitted by 20-Mar-2023.  If it is not possible for you to submit your revision by this date, we may have to consider your paper as a new submission.  
  
Important note: After a manuscript has been accepted for publication, the author’s company or institution will be requested to pay an optional $110 per page to cover part of the cost of publication. These page charges are not obligatory, nor is their payment a prerequisite for publication as long as the main paper does not exceed the page limit (10 pages for a regular paper, 15 pages for a survey paper, and 6 pages for a letter). Papers that exceed this page limit will be requested to pay a mandatory over-length page charge of $200 per page, for every page over these limits. The maximum manuscript length (excluding supplementary materials) with over-length page charge is 15 pages for a regular paper, 21 pages for a survey paper, and 9 pages for a letter. Supplementary materials are not included in determining over-length pages, and get published online with the manuscript. Authors are strongly advised to read the up-to-date instructions at <https://nam04.safelinks.protection.outlook.com/?url=https%3A%2F%2Fcis.ieee.org%2Fpublications%2Fieee-transactions-on-artificial-intelligence%2Finformation-for-authors-tai&data=05%7C01%7Cnsh%40usf.edu%7Ca17ae668713c4a0b6f1c08db119a6c2d%7C741bf7dee2e546df8d6782607df9deaa%7C0%7C0%7C638123127175085062%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=x9kKM5DhTjeAhJu5XRg5tPYp5XJ1GrTHdfFmYY%2FsjSI%3D&reserved=0>  
  
Once again, thank you for submitting your manuscript to the IEEE Transactions on Artificial Intelligence and I look forward to receiving your revision.  
  
Sincerely,  
  
Prof. Hussein Abbass, Fellow IEEE  
Editor-in-Chief, IEEE Transactions on Artificial Intelligence  
ieee.tai.eic@gmail.com, h.abbass@unsw.edu.au  
  
  
Editor-in-Chief Comments to Author:  
  
Editor-in-Chief  
  
Please remove abbreviations from the title. "SIMPOR" does not contribute information to the title; please remove.  
  
Associate Editor Comments to Author:  
  
Associate Editor  
Comments to the Author:  
Please follow reviewers' comments to further revise the paper in better quality.  
  
  
  
  
Reviewer(s)' Comments to Author:  
  
Reviewer: 1  
  
Comments to the Author  
This work explores how class imbalanced data  
affects deep learning and proposes a data balancing technique  
for mitigation by generating more synthetic data for the minority  
class. The paper is well-written and easy to read. Several comments are listed below.  
  
1. The detail of entropy-based active learning should be improved. For example, how many samples are used for model training initially? On p.3, the classifier with \theta^0 is trained using a random batch of k selected samples. So only k samples are randomly drawn from the training set to train the model? Besides, the authors further state that the informative samples are obtained by selecting k samples based on the top k highest entropy. These two k are identical?  
2. The authors simplify Eq. (8) by assume that the prior probabilities in the two classes are identical to cancel out this term. I can understand this assumption can make it easy to calculate, but is it reasonable? The priors for the majority and minority classes are apparently different. I suggest the authors consider the empirical priors to conduct experiments.  
3. The proposed method involves many hyper-parameters. For example, h in Eq. (12). The values of these hyper-parameters and how to obtain them should be presented in the manuscript.  
4. The comparison methods should include recently published methods and deep learning methods. Among the comparison methods, only GOD was a method published in 2022. SMOTE, Boderline-SMOTE, EE, and ADASYN where published before 2010. SVMCS was in 2012.  
5. In Section VI. E (Data Visualization), how to calculate No. of intersection samples?  
6. In Fig. 6, how to obtain the densities of the results? If they are obtained by using KDE, are they unfair to the other methods that don't use KDE?  
7. In Fig. 7 & 8, the results show that \alpha and IP do not significantly affect the performance, but they are counter-intuitive. The authors should provide more discussion to explain the results rather than only presenting the results.  
8. In Section VI. F (Processing Time), the sentence "To explore more ... the To better evaluate the technique" should be re-phrased.  
  
Reviewer: 2  
  
Comments to the Author  
This work explores how class-imbalanced data affects deep learning and proposes a data-balancing technique to mitigate by generating more synthetic data for minority classes.  
The author has done a lot of work, and the experimental results are very good. The content of the article is substantial. I hope the following comments could be useful for the further improvement of the paper:  
1. The language of the manuscript is not academic enough. There are too many "we" in the text. I suggest the writer use the passive voice  
2. The contribution: 3) We applied our technique to 41 real datasets with a diversity of imbalance ratio and the number of features. I don't think this is a contribution. This is a statement of work.  
3. AUC is an acronym. It is better to use an acronym after mentioning the full name in the “Introduction”.  
4. The letters used in the formulas are not explained clearly, especially formula (9).  
5. More new published methods are suggested to be used as the comparison algorithms.  
6."Related work" should be put to the second section of the paper. In addition, the following references are suggested to be present and cited:  
[1]Wei Feng\*, Gabriel Dauphin, Wenjiang Huang, Yinghui Quan, Wenxin  g Bao, Mingquan Wu, Qiang Li, “Dynamic synthetic minority over-sampling technique based rotation forest for the classification of imbal a nced hyperspectral data,” IEEE Journal of Selected Topics in AppliedEarth Observations and Remote Sensing, vol. 12, no. 7, pp. 2159–2169, 2019.  
[2]Wei Feng\*, Wenjiang Huang and Wenxing Bao, "Imbal an ced Hyperspectral Image Classification With an Adaptive Ensemble Method Based on SMOTE and Rotatio n F orest With Differentiated Sampling Rates,"  IEEE Geoscience and Remote Sensing Le tters, vol. 16, no. 12, pare suggestedggestedc. 2019.  
[EXTERNAL EMAIL] DO NOT CLICK links or attachments unless you recognize the sender and know the content is safe.