Author’s response to reviews

Title: Risk Factor Analysis of Device-related Infections: Value of Re-sampling Method on the Real-World Imbalanced Dataset

Authors:

Xiang-Fei Feng (fengxiangfei@xinhuamed.com.cn)
Ling-Chao Yang (yanglingchao@xinhuamed.com.cn)
Li-Zhuang Tan (lzhtan@bjtu.edu.cn)
Yi-Gang Li (drliyigang@yeah.net)

Version: 3 Date: 16 Jul 2019

Author’s response to reviews:

Carlos Luis Sánchez Bocanegra (Reviewer 1): Dear all:

First of all, congratulations for this study. taking in consideration the duration for the study it could be a great effort for recollecting those information. Hope you can find a good clue for a excellence draft.

Let me introduce the considerations as a major revision for this proposal:

1. Introduction: I think it needs a re-write. My reading of the state of the art is complex and heavy. I should recommend a simply description of this section for example "Some biomedical data belong .....subsets[10]" feel a little boring piece of text. think about 1 paragraph 1 idea, 1 sentece describe de simple esence of the resolution/problem.

Authors’ response: Thank you very much for your advice. We have re-written the part of “Introduction”, and deleted the part of "Some biomedical data .....subsets[10]".

2. Method: I feel loosing, where do you define this section? As an example: in result the draft expose the re-sampling (do you think this could be a method section?). Please be as clear as possible and try not to mix section to be simply and transparent.
Authors’ response: Thank you very much for your advice. We have re-written the part of “Definitions” and deleted the overlap section about “re-sampling” in result. We added the part of “Accuracy rate testing”.

3. Evaluation: I think you take a good approach, simple and easy to read. Perhaps I should explain better ROC graph.

Authors’ response: Thank you very much for your advice. We have added ROC graph after resampling and explained better ROC graph. We added the part of “Accuracy rate results” in “Results” and “Evaluation” in “Discussion”.

4. Bibliography: I think you need to update this revision of the bibliography. References from 1994, 2001 should be taken a long time?

Authors’ response: Thank you very much for your advice. We have updated this revision of the bibliography, and deleted the references before 2001, added the latest literatures.

Hansi Zhang, MS (Reviewer 2): This paper talks about author's experience on resampling of the imbalanced Device-related Infections (FRI) dataset. The study is well-motivated and the result is well-presented. The authors has made a lot of efforts on revisions. I only have a few small comments as listed below:

(1) The methods used in this paper to re-sample the imbalanced data is presented. However, as mentioned that imbalanced data problem occurs frequently in biomedical data. How others people deal with imbalanced data? What method do they use? What's the benefits of using the study method? More discussion on related work is needed.

Authors’ response: Thank you very much for your advice. We have added the section “Methods for addressing imbalance” about imbalanced data problem in discussion. We also added pie plot about CIEDI difference between male and female. We added the part of “Accuracy rate results” in “Results” and “Evaluation” in “Discussion”.

(2) Page 12 line 34 to page 13 ling 9. I wonder if the description of re-sampling would fit better in methods than discussion?

Authors’ response: Thank you very much for your advice. We have moved the section “re-sampling” in discussion to the section “re-sampling” in methods.
(3) When discussing the limitation of the work, what's the possible solutions for the further work?

Authors’ response: Thank you very much for your advice. We have added the section “Evaluation” in discussion, and added the possible solutions for the further work in “Study limitations”. √