26-Feb-2019  
  
RE: TCBB-2018-05-0199.R1, "MGRFE: multilayer recursive feature elimination based on an embedded genetic algorithm for cancer classification"  
Manuscript Type: Regular Paper  
  
Dear Dr. Li,  
  
We have completed the review process of the above referenced paper for the IEEE/ACM Transactions on Computational Biology and Bioinformatics and recommend that your paper undergo a Major Revision.  
  
Enclosed are your reviews. If you should choose to revise your paper, please prepare a separate document describing in detail how each of the reviewers' comments are responded to in your revision and submit it before 27-May-2019.  
  
To revise your manuscript, log into <https://mc.manuscriptcentral.com/tcbb-cs> 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.  
  
Once the revised manuscript is prepared, you can upload it and submit it through your Author Center.  
  
Text in any color other than black is not acceptable. Your revised paper must include the following:  
  
-abstract  
-index terms  
-author affiliation information  
-main text  
-references  
-figure captions  
-table titles  
-brief author biographies  
(biographies are not required for short papers or comments papers)  
  
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)’ questions and comments. You may also upload your responses as separate files for review along with your revision. If you choose to do this, please choose “Summary of Changes” as the file designation.  
  
When the submission process is complete, you will receive an automated confirmation email immediately. If you did not receive that email, your submission is not yet complete.  
  
The journal’s Administrator will contact you should we have any concerns or questions regarding your revision. Otherwise, your revision will be forwarded to the assigned Associate Editor to begin a second round of reviews.  
  
Our page limitation and formatting guidelines for TCBB can be found on:  
  
<http://www.computer.org/portal/web/peerreviewjournals/author#manuscript>  
  
Thank you for your contribution to TCBB, and we look forward to receiving your revised manuscript.  
  
Sincerely,  
  
Aidong Zhang, EIC  
IEEE/ACM Transactions on Computational Biology and Bioinformatics  
[aidong@virginia.edu](mailto:aidong@virginia.edu)  
  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
Editor Comments  
  
Associate Editor  
Comments to the Author:  
The manuscript was reviewed by the original reviewers.  
Although Reviewer 1 is satisfied with the revised version, Reviewer 2 gives very critical comments.  
  
Therefore, I recommend the authors to revise the manuscript with taking all comments into account.  
  
Since I understand that giving theoretical justification is difficult,  
it is enough to give some discussions.  
  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
  
Reviewer: 1  
  
Recommendation: Accept With No Changes  
  
Comments:  
The revisions are satisfactory.  
  
Additional Questions:  
1. Which category describes this manuscript?: Research/Technology  
  
2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under Public Comments below.: Relevant  
  
1. Please explain how this manuscript advances this field of research and/or contributes something new to the literature.: This research proposed a new recursive gene elimination technique which further enhance the research of gene selection towards the fight against carcinogenic diseases.  
  
2. Is the manuscript technically sound? Please explain your answer under Public Comments below.: Yes  
  
1. Are the title, abstract, and keywords appropriate? Please explain under Public Comments below.: Yes  
  
2. Does the manuscript contain sufficient and appropriate references? Please explain under Public Comments below.: References are sufficient and appropriate  
  
3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under Public Comments below.: Could be improved  
  
4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain under Public Comments below.: Could be improved  
  
5. Please rate the readability of the manuscript. Explain your rating under Public Comments below.: Readable - but requires some effort to understand  
  
6. Should the supplemental material be included? (Click on the Supplementary Files icon to view files): Yes, as part of the main paper if accepted (cannot exceed the strict page limit)  
  
7. If yes to 6, should it be accepted: After revisions. Please include explanation under Public Comments below.  
  
Please rate the manuscript. Please explain your answer.: Good

Reviewer: 2  
  
Recommendation: Reject  
  
Comments:  
The selection of highly informative genes in cancer patients is a standard problem with many techniques in existence. The paper presents yet another approach based on an embedded genetic algorithm. In my previous review I had raised a number of queries, which have essentially been dismissed by the authors in their revised version. My queries have NOT been addressed satisfactorily.  
  
My original comment 3 is that there is no validation on an independent data set. The authors state in their rebuttal that "Thus, the currently published gene selection algorithms on microarrays are commonly validated within each microarray benchmark dataset."  
  
I am sorry to say that this is incorrect. I have published several papers in computational cancer biology, and ALL of them had validations on independent data sets. I am not persuaded by the authors' argument.  
  
"For microarray benchmark datasets about same disease, the features and sample classes are usually different. Different microarray datasets usually have different gene features for the gene probes vary among different microarray analysis platform. For example, on the leukemia related datasets of Leuk and MLL used in this study, the gene probes are very different for generating from different microarray platforms."  
  
This is PRECISELY the reason why validation on an independent data set is so crucial. It is true that two different databases of the same form of cancer may have different genes under study. The way to handle this is to study only those genes that are common to both databases. One can also convert microarray values to Z-scores by subtracting the sample mean and dividing by the sample variance. The authors don't even try to do this.  
  
"Thus, the currently published gene selection algorithms on microarrays are commonly validated within each microarray benchmark dataset."  
  
This is not correct. The authors are simply trying to justify whey they did not do any validation on an independent dataset.  
  
If they have managed to do cross-validation on another dataset for leukemia, then that should be in the main paper, not in the supplementary material.  
  
My comment 4 was that their method lacked theoretical justification and compared it to SVM-RFE. Here again the authors simply explain away my objection. They say that their GA (genetic algorithm) works faster than that of Kar et al. That was not my point at all.  
  
In short, I believe that the authors have not adequately addressed my previous comments. Without either theoretical justification or validation on independent datasets, there is very little merit in the paper.  
  
Additional Questions:  
1. Which category describes this manuscript?: Practice / Application / Case Study / Experience Report  
  
2. How relevant is this manuscript to the readers of this periodical? Please explain your rating under Public Comments below.: Relevant  
  
1. Please explain how this manuscript advances this field of research and/or contributes something new to the literature.: The selection of highly informative genes in cancer patients is a standard problem with many techniques in existence. The paper presents yet another approach based on an embedded genetic algorithm. In my previous review I had raised a number of queries, which have essentially been dismissed by the authors in their revised version. My queries have NOT been addressed satisfactorily.  
  
2. Is the manuscript technically sound? Please explain your answer under Public Comments below.: Partially  
  
1. Are the title, abstract, and keywords appropriate? Please explain under Public Comments below.: Yes  
  
2. Does the manuscript contain sufficient and appropriate references? Please explain under Public Comments below.: References are sufficient and appropriate  
  
3. Does the introduction state the objectives of the manuscript in terms that encourage the reader to read on? Please explain your answer under Public Comments below.: Could be improved  
  
4. How would you rate the organization of the manuscript? Is it focused? Is the length appropriate for the topic? Please explain under Public Comments below.: Could be improved  
  
5. Please rate the readability of the manuscript. Explain your rating under Public Comments below.: Easy to read  
  
6. Should the supplemental material be included? (Click on the Supplementary Files icon to view files): Yes, as part of the digital library for this submission if accepted  
  
7. If yes to 6, should it be accepted: As is  
  
Please rate the manuscript. Please explain your answer.: Poor