OPEN PEER REVIEW

Open Peer Review (OPR)

As a new feature for progressing towards transparency, we decided to open a new window for all of our editors as well as authors titled "**Open Peer Review**". We hope by this new facility, our reviewers will be more motivated and authors will be more satisfied with the review process. We believe that publishing our peer review reports could make a transparent and clear environment for all our efforts within a journal, but not all reviewers tend to publish their comments.

What is "Open Peer Review" process?

An "Open Peer Review" process is making the details of all review process (including reviewers, associate editors, and EICs comments) as "Public" as it is agreed by EIC, Authors, and reviewers.

Advantages of "Open Peer Review" process

- More transparency, constructiveness and tactful comments of the peer review process: leads to an increase in the quality of reviews
- More motivations for all involved roles in the review process
- Authors' satisfactions from the review process: Increases honesty between authors and reviewers
- Education of both authors and new students
- Prevents reviewers from following their individual agendas and leading to the detection of reviewers' conflicts of interests

You can find out more at:
https://brieflands.com/briefland/knowledgebase/category/tree.html#opr.html

Iranian Journal of Pediatrics
Article DOI: 10.5812/ijp.103094
Published in: Iranian Journal of Pediatrics: 31(2); e103094

Peer Review Report for "Designing a Data Mining System to Predict Treatment-Requiring Retinopathy of Prematurity in Neonates: A Pilot Study"

Author(s): Farshid Khorasani , Ramak Roohi poor, Afsar Dastjani Farahani , Azam Orooji , Mohammad Reza Zarkesh

Review Timeline:

Submit Date:	29 Mar 2020
Revised Date:	14 Jan 2021
Accepted Date:	17 Jan 2021

Revision (0)

Here, you can see the **Reviewers**, **Associate Editors** and **EICs'** comments from the beginning to the end of the revision process.

REFEREE: Reviewer | **Revision** (0) 31 Dec 2020

Dear AE,

- Data mining is the process of discovering patterns (i.e. knowledge discovery) in large data sets, while the methods used in the paper are considered training classifiers using limited data set; Although they both use methods like machine learning, but their goal are different. So the title of the paper is not relevant.

- Any classifier trained using machine learning techniques over small and not-so-diverse data tends to overfit the training set (i.e. working perfect on training set and resulting poor on unseen data). This is specially affecting the validity of Naive Bayes classifier due to the assumption used: that is "Bayes' theorem with strong independence" which assumes that input factors are completely independent from each other, which is definitely not the case here.

Validity:

- The data set used for training classifiers is small in size, even related to those mentioned on referenced articles. This is also mentioned at Discussion section.

- Some of the risk factors like "intubation period" and "oxygenation period" seem

problematic to be used as decision factors;
- These are actually measured after the treatment process is done, i.e. these factors are not available when deciding on ROP treatment-requirement.
- It is not clear what is the amount of missing data on risk factors, and what has been done to mitigate the effect of lack of these data. No details on data preprocessing is given in general.
Linguistics:
- Introduction Section: "Recently, data mining methods as valuable techniques are implemented in the health care system"; they are not usually "implemented" in health care systems.
- Materials and Methods Section, Primary Outcomes: "Our primary objective was to suggest some criteria for predicting treatment-requiring retinopathy of prematurity in neonates."; This is only meaningful for decision tree classifier, while that classifier is neither the main goal nor the selected classifier of the paper.
- Results Section: "Results showed that the sensitivities correlated with Random Forest and Naive Bayes models were higher in comparison with other techniques."; the word "correlated" is not relevant here. The preposition "of" should be used instead.
- Mentions of "performance metrics accuracy, precision, sensitivity, specificity, and F-measure" are repeated in many paragraphs, sometimes with no analysis.
Thanks

REFEREE: Associate Editor | **Revision** (0) 1 Jan 2021

Dear EIC,
The reviewers and editors were very interested in your work, which they found timely, novel, and important. They however had a few minor concerns regarding your manuscript, as detailed in their attached reviews.
Addressing these concerns will require a revision of your manuscript before it can be reconsidered for publication in the Journal.
Thanks

REFEREE: EIC | **Revision** (0) 1 Jan 2021

Dear Authors,
Thank you for submitting your manuscript to Iran J Pediatr. Your manuscript has now been reviewed by a reviewer and an Associate Editor. I now have the reports of the Referees, which suggest that your manuscript might be suitable for publication following a number of revisions. I attach some reviewer comments and invite you to submit a revised version of your manuscript addressing the points made.
Thanks

OPEN PEER REVIEW

Revision (1)

Reply to Reviewers

Ideally, the reviewing process can significantly improve the submitted manuscripts by allowing the authors to take into account the advice of reviewers. Author(s) must reply to all reviewers' comments in a separate Word file, point by point. A "**Reply to Reviewers**" document is submitted along with revised manuscript during submission of revised files, summarizing the changes that the authors made in response to the reviewers' comments. The responses to reviewers' comments specifies how the authors addressed each comment the reviewers made.

You can read the authors' responses to the reviewers' comments in the next page.

Reviewer 1:
Technical:

Data mining is the process of discovering patterns (i.e. knowledge discovery) in large data sets, while the methods used in the paper are considered training classifiers using limited data set; Although they both use methods like machine learning, but their goal are different. So the title of the paper is not relevant.

New title: Designing a data mining system to predict treatment-requiring retinopathy of prematurity in neonates: A pilot study

Although machine learning algorithms require a huge number of samples for training (1,2). However, collecting large number of samples is not always feasible. This is especially true in the medical field, when we are dealing with data on a rare disease or when, for whatever reason, limited samples are available (3,4).

Based on the results of a meta-analysis study that systematically reviewed 42 articles including 18,000 premature infants, the prevalence of ROP was reported to be 23.5% (95% CI: 20.4–26.8) in Iran. The prevalence of ROP stages 1, 2, 3, 4 and 5 was 7.9% (95% CI: 5.3–11.5), 9.7% (95% CI: 6.1–15.3), 2.8% (95% CI: 1.6–4.9), 2.9% (95% CI: 1.9–4.5) and 3.6% (95% CI: 2.4–5.2), respectively (3).

Any classifier trained using machine learning techniques over small and not-so-diverse data tends to overfit the training set (i.e. working perfect on training set and resulting poor on unseen data). This is specially affecting the validity of Naive Bayes classifier due to the assumption used: that is "Bayes' theorem with strong independence" which assumes that input factors are completely independent from each other, which is definitely not the case here.

Although the sample size is small, but samples are selected completely randomly therefore, there is no bias.

The performances of data mining algorithms are depended on the nature of the data. Hence, we choose the best algorithm by implementing different algorithms (e.g. tree based, Function based, Meta and Bayesian), considering comprehensive set of performance criteria and evaluation based on 10-fold cross validation.

Previous studies have shown that naïve Bayes is suitable for a small dataset with high correlation between the task and other non-task attribute variables (6).

Moreover, SVM is also suitable for the balanced dataset with the binary-class task (6). The results of applying SVM to our dataset are shown below, which are not better than the results obtained.

Accuracy: 77.1429

Sensitivity: 65.7

Specificity: 88.6

F-Measure: 74.2

Precision: 85.2

Validity:

The data set used for training classifiers is small in size, even related to those mentioned on referenced articles. This is also mentioned at Discussion section.

Some of the risk factors like "intubation period" and "oxygenation period" seem problematic to be used as decision factors; These are actually measured after the treatment process is done, i.e. these factors are not available when deciding on ROP treatment-requirement.

It is not clear what is the amount of missing data on risk factors, and what has been done to mitigate the effect of lack of these data. No details on data preprocessing is given in general.

Data are collected based on a “data extraction form”. Then, six records with more than 70% missing data, were excluded from analysis. There were 17 unregistered data elements that were replaced using mean and median for continuous and discrete attributes, respectively.

References:

1. Majumdar A, Singh R, Vatsa M. Face verification via class sparsity based supervised encoding. IEEE transactions on pattern analysis and machine intelligence. 2016 May 17;39(6):1273-80.https://doi.org/10.1109/TPAMI.2016.2569436
2. Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. InProceedings of the IEEE conference on computer vision and pattern recognition 2017 (pp. 4700-4708).
3. Thomas RM, Bruin W, Zhutovsky P, van Wingen G. Dealing with missing data, small sample sizes, and heterogeneity in machine learning studies of brain disorders. InMachine Learning 2020 Jan 1 (pp. 249-266). Academic Press.https://doi.org/10.1016/B978-0-12-815739-8.00014-6
4. Keshari R, Ghosh S, Chhabra S, Vatsa M, Singh R. Unravelling Small Sample Size Problems in the Deep Learning World. In2020 IEEE Sixth International Conference on Multimedia Big Data (BigMM) 2020 Sep 24 (pp. 134-143). IEEE.
5. Azami M, Jaafari Z, Rahmati S, Farahani AD, Badfar G. Prevalence and risk factors of retinopathy of prematurity in Iran: a systematic review and meta-analysis. BMC ophthalmology. 2018 Dec 1;18(1):83.https://doi.org/10.1186/s12886-018-0732-3
6. Zhang Y, Xin Y, Li Q, Ma J, Li S, Lv X, Lv W. Empirical study of seven data mining algorithms on different characteristics of datasets for biomedical classification applications. Biomedical engineering online. 2017 Dec 1;16(1):125.https://doi.org/10.1186/s12938-017-0416-x

OPEN PEER REVIEW

Revision (1)

Here, you can see the **Reviewers**, **Associate Editors** and **EICs'** comments from the beginning to the end of the revision process.

REFEREE: Associate Editor | **Revision** (1) 17 Jan 2021

Dear EIC,
As the authors improved their manuscript based on the reviewer's recommendations, I think it can be accepted.
Thanks

REFEREE: EIC | **Revision** (1) 17 Jan 2021

Dear Authors,
The second round of peer review of your manuscript has been completed, and I am pleased to inform you that your manuscript has now been accepted for publication in our journal.
Thanks