

Truyen Tran

From: onbehalfof+jbhi-eic+embs.org@manuscriptcentral.com on behalf of jbhi-eic@embs.org
Sent: Tuesday, 25 February 2014 10:47 AM
To: Truyen Tran
Subject: Major Revision - TITB-00720-2013-JBHI

24-Feb-2014

JBHI Ref: TITB-00720-2013
Reject/Resubmit (major revision and new external review required)

Dear Dr. Truyen Tran,

This letter is to inform you that the peer review process has concluded for manuscript, "Predicting Unplanned Readmission for Heart Failure from Electronic Medical Records: Stabilizing High-Dimensional Models Using Feature Graphs," which you had submitted for possible publication in the IEEE Journal of Biomedical and Health Informatics (J-BHI).

The Associate Editor responsible for your manuscript review has received feedback from independent reviewers and compiled their evaluations. It is the recommendation of the Associate Editor and the Editor-in-Chief that your manuscript requires a MAJOR REVISION before it can be accepted for publication in J-BHI. Please note the J-BHI Editorial Policy that only one major revision is allowed for any submitted manuscript. This means that if the review recommendation for your revised paper triggers another major revision, the paper will be rejected automatically.

Enclosed, please find the comments by the Associate Editor and all the reviewers. I hope that the feedback is helpful for further improving the quality of your manuscript. If you decide to resubmit a revised manuscript, this must be done within 10 weeks (not extendable) from the date of this message. Please quote the above manuscript reference number for all future correspondence.

Submission of the revised manuscript should be made via Manuscript Central, where it is now possible to upload a revised version of your paper. Please do not submit your revision via email, regular mail, or fax, as they will not be considered for publication.

Also note that it is mandatory to enter your replies to the reviewers' questions and indicate how you have dealt with their comments in the revised manuscript. Please include your replies in the authors' response section or a separate file with a point-by-point explanation of all the changes made. Please do NOT include it in the cover letter since this is not accessible by the reviewers.

Please also note the J-BHI editorial policy as detailed at <http://jbhi.embs.org/for-authors/editorial-policy/> and carefully read about how to prepare your manuscript (<http://jbhi.embs.org/for-authors/prepare-and-submit-your-manuscript/>). You should also note the mandatory charge for over-length papers as imposed by IEEE (the nominal length for a regular paper is 7 pages when formatted in IEEE double column style).

Thank you very much for considering JBHI to publish your research work.

Sincerely,

Prof. Guang-Zhong Yang
Editor-in-Chief

Cc: file

Associate Editor's comments to the authors:

Associate Editor

Comments to the Author:

(There are no comments. Please check to see if comments were included as a file attachment with this e-mail or as an attachment in your Author Center.)

Reviewers' comments to the authors:

Reviewer: 1

Comments to the Corresponding Author

The authors developed a model to predict readmission of heart failure patients. The model was trained using patients' data and which were extracted from EHR. To improve the stability of the trained model, the authors designed a feature graph using ICD10 relationships and temporal constraints. The results show promising improvements.

For table 2, I would recommend adding classification precision and recall rate, which can improve the paper's readability.

The author proposed a method for stabilizing the model using feature graphs. I am curious to know how the performance changes with and without the stabilization.

I would recommend the authors testing cross-validation for evaluating the stability of the model which can be used to compare with the Laplacian regularization.

Overall, the paper is well-written and the results are interesting.

Reviewer: 2

Comments to the Corresponding Author

Summary

The paper describes an approach to building models for prediction of readmission for heart failure. The model is constructed from the EMR data. The problem of data dimensionality is tackled by feature selection using lasso and by another regulation term that works towards assigning similar weights to related features. The feature graph is constructed by connecting the pairs of feature which differ only in their time index or in a part of their ICD code.

The main claims of the paper is that the resulting models are comparable in quality with the existing models and that they are more stable with regard to the selection of features.

The idea of using the feature graphs to regularize the logistic regression is nice -- although not entirely original. Another nice idea of the paper is the simple construction of the feature graph.

On the downside, the paper may lack enough new contributions and a more thorough experimental validation that would show the advantage of using the feature graphs in the presented context.

A general comment

It is unclear what is the main contribution of the paper.

- Based on the title, the new idea is the use of feature graphs (applied to prediction of heart failure-related readmission). The idea of using feature graphs is not original (another interesting recent example is Cun & Froehlich:

Network and Data Integration for Biomarker Signature Discovery via Network Smoothed T-Statistics, PLOS One, 2013, which uses Laplacean for a similar purpose). As the paper itself states, these have already been used in genetics for quite some time. Solely the fact that this may be their first use in health informatics does not represent a significant scientific advance. Furthermore, the paper does not compare the resulting model with a model constructed from the same data but without the Laplacean.

- Another interpretation of the title, with relation to the contribution, could be that the paper presents a new model. The new model is, however, at most comparable with the existing models in terms of accuracy, and it offers no significant advantages.

- Better stability of the resulting models and the use of consistency and the Jaccard index to assess it: this aspect is not put at the front and not explored in much detail. It is also not clear why a greater stability would be desired. In the bias-variance framework, the stability is related to the variance which represents a part of the predictive error. Here, however, the stability in question is not the stability of the models results but of the selected features, which does not necessarily impact the models accuracy.

Particular comments

Originality: In section II.e, it should be made clear that the content of this section is not original but comes from [20] (or, perhaps, an earlier work). Without stating this explicitly, the reader could have an impression that this is a new scientific contribution.

This issue is further muddled with, for instance, "Our approach consists of a novel technique to mitigate the problem by utilizing feature graphs (...)".

Results: the paper does not compare the models with and without the regularization this same data. It is therefore impossible to judge whether the regularization helps.

The bulk of the experimental part is related to exploring the AUCs at different values of coefficients. However, these differences are minor and within the confidence intervals. The appropriate conclusion from them would be that the value of the regularization coefficients have no effect on the AUC. The highest AUCs are achieved at the minimal values (in Table II), so it would seem to make sense the even lower values (including, in particular, 0 -- no normalization).

The differences in Figure 5 are exaggerated due to misleading scale y. The scale should show the entire AUC, from 0.5 to 1 (assuming that the classifier is better than random). I am attaching a file with such a figure, which shows that the regularization coefficient do not matter.

Minor comments

Section III. b: What is meant by "model generalizability"?

Figure 4: There is no need to put arrows on the lines since the graph is undirected.

Section IV. a, first sentences: Why would this be an advantage? What is wrong with manually constructed, hypothesis-based models, especially if they are more accurate?

Moreover, data-driven and automatically constructed models are advantage over ... which models? Logistic regression models (with or without regularization) are always data-driven and automatically constructed, so the presented technique does not present an advantage in this respect.

Similarly, the sentence that the "additional data may improve the accuracy" is a common expectation; on the other side, the paper does not present any evidence for this.

Does sparsity suggest a better interpretation? Or are models with less variables just easier to interpret?

Section IV. a, last paragraph: I would not agree that the discrimination power suggest that the model could be used in practice. An AUC of 0.66 means that the model is able to distinguish between a pair of patients with different outcomes in two cases out of three, which is rather low (although not worse than the existing models).

Please check for grammatical errors. The paper is generally written well, yet there are some possible mistakes.

the weights of weak features are derived (driven?) towards zero derivation and validation cohorts were separately (separated?) reaches the pick (peak?) Reference 29: mann-whitney -> Mann-Whitney Reference 31: jaccard -> Jaccard

Reviewer: 3

Comments to the Corresponding Author

Dear authors,

As we all know medical records exist high dimensional, sparse and small samples. There are always many missing parts in the sample records, and my question is how to deal with them? To discard or other ways? My second problem about your paper is that it's not very clear to tell the predicting model. That is I cannot very clearly understand what you method is and how innovative you method is in your abstract. So maybe you can improve it and make it clear. The end question is that you method's accuracy is about 60%, I'd like to know whether you can compare your way with other machine learning methods, such as NN or others.

In the end I think it's meaningful to predict unplanned readmission for heart failure from EMR. It's also very significant to deal with medical records although its result such as accuracy is not very high or its method is not very new compared with machine learning. So dear authors, my opinion is the paper can be accepted if you are willing to improve it.

OVERVIEW OF CHANGES

Article Title	Stabilizing High-Dimensional Models using Feature Graphs
Submitted to	IEEE Journal of Biomedical and Health Informatics (J-BHI)
Ref. No.	TITB-00720-2013.R1
Authors	Shivapratap Gopakumar, Truyen Tran, Tu Dinh Nguyen, Dinh Phung, Svetha Venkatesh
Corresponding contact	Dr Truyen Tran Center for Pattern Recognition and Data Analytics, Deakin University, Australia. Email: truyen.tran@deakin.edu.au Tel: +61 3 5227 2082, Fax: +61 3 5227 2028
Last updated	July 15, 2014

Overview

We thank the editor and the reviewer for the opportunity to improve our work. We have carefully considered the reviewer's concern. In the revised manuscript, we have added predictive performance analysis of our proposed stabilization method.

Reviewer #1

Comments:

1. *The revised manuscript did not address one of my major concern, that is how the stabilized model leads to better predictive performance, such as reducing errors and improving accuracy. It is very important for assessing the contribution of the proposed method.*
 - (a) We thank the reviewer for pointing out the limitation in our work. In accordance with the reviewer's views, we have evaluated our stabilized model on the basis of classification accuracy, ROC curve analysis and goodness-of-fit. Accordingly, the following changes have been made to the paper:
 - i. In **Section II, D (Validation Protocol)**, we have added a subsection on measuring goodness-of-fit using Hosmer-Lemeshow test.
 - ii. In **Section III (Results)**, we have re-examined our model with the validation cohort for overall classification accuracy. Further, we perform ROC curve analysis and measure the goodness-of-fit for the unstable and stabilized models. The summary of Hosmer-Lemeshow goodness-of-fit test has been included.
 - iii. In **Section IV. (Discussion and Conclusion)**, we have added the implications of the results obtained.
 - iv. The application of Laplacian graph regularization significantly improves the goodness-of-fit of the model and marginally improves the classification accuracy. We have updated the abstract and introduction section to reflect these new results.