Reviewer’s report

Title: A classification framework for exploiting sparse multi-variate temporal features with application to adverse drug event detection in medical records

Version: 1  Date: 11 Oct 2018

Reviewer: Cynthia Sung

Reviewer's report:

The main innovation of this paper appears to be a novel way of representing a temporal sequence of events when there is a high proportion of sparsity. The paper clarifies several points over the previous version. The abstract states that "A large-scale evaluation on 15 ADE datasets extracted from a real-world EHR system shows that the proposed framework achieves significantly improved predictive performance compared to a competitor approach" The 15 ADEs selected are less than 10% of the ADE in categories A1 and A2 of the referenced paper, so this is really a very small sample. A more comprehensive test of all or at least a much larger subset of the ADEs in A1 and A2 would add to the robustness, and since it is computational work, can the authors crunch those numbers? While a pairwise comparison with sl or plain, lr yielded significantly more information content, the authors did not really test the *predictive* capability of their method for identifying an ADE and then check how often the prediction was true from the diagnostic codes of encounters that had not been used in the initial round of method development. The finding needs to be modified with more modest and precise language what actually was demonstrated.

From a clinical perspective, it's difficult to appreciate the value of the kind of information presented in Table 4, how it will be used to find ADEs. Not being an expert in the mathematical formalisms, I am unable to review this aspect of the work. As possibly one small step towards a bigger effort to extract more information from data with a lot of sparsity and uneven temporal spacing of information, this article may be of interest to algorithm developers, but the authors should not overstate what has been demonstrated.

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

No

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Yes
Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

No

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I recommend additional statistical review

Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons
CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal