Reviewer’s report

Title: Interpreting Patient-Specific Risk Prediction Using Contextual Decomposition of BiLSTMs: Application to Children with Asthma

Version: 0 Date: 26 Jun 2019

Reviewer: Primoz Kocbek

Reviewer's report:

The study adapts and extends contextual decomposition (CD), an algorithm for producing interpretable phrase-level importance scores, in the context of predicting future clinical outcomes using EHR data. More precisely, it analyses the predictions of which pre-school children with respiratory system related complications will have asthma at school-age, where it focuses patient visits with a higher impact on the predicted outcome from EHR, ie. it identifies and validates the relevant subset of visits driving the predictions. The study further presents quantitative (validation of CD via additional artificial visits, correlation of three highest CD scores with logistic regression coefficients) and qualitative evidence (explaining prediction using individual visit score, using relative contributions of subsets of visits, identifying top scoring patterns).

The study provides an interesting approach to the interpretability of DL models (LSTM and BiLSTM) in the context of longitudinal EHR data and presented quantitative and qualitative evidence to explain patient-specific predictions, which might be useful to physicians.

I do have some comments, that might need addressing:

-the approach described uses a sliding window approach to find the most predictive subset of visits, as the authors noted this is a greedy algorithm and usually the number of visits is low (it was defined as less then 10), but it would make to sense to suggest/propose an alternative approach in cases where the number of visits is not low, eg. a longer time period and patients with a chronic condition (toward the end of the manuscript, they describe examples with 19 and 14 visits);

-the authors choose a scenario where the proportion of cases is high (~55%), that might not be a general scenario, which might be a less balanced scenario for prediction and it should be noted;

- results were presented as average AUC values values for 10 iterations and in addition to the mean one would expect that there were some additional information provided about the AUC, maybe confidence intervals, so the reader might get a sense of the variability of the AUC.
Minor:

- page 6, line 47: the word "that" appears twice consequently

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

Yes

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.

Yes

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 am able to assess the statistics

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