Reviewer’s report

Title: Predicting life expectancy with a long short-term memory recurrent neural network using electronic medical records

Version: 1 Date: 19 Jul 2018

Reviewer: Shahryar Eivazzadeh

Reviewer's report:

Summary:

The study presents a model/method for the prognoses of death of patients using their EMRs. The study is based on using LSTM models, and uses both structured and unstructured (text) data. The study was able to outperform doctors' prognoses. However, the model cannot specify determinant features in prognosis due to the nature of NN.

General/Major:

1- Improving ACP and palliative care can be an output to using your models and methods but as far as I can see your study is not limited to this context and the results can have other applications (like better procurement of care facilities). But, the wording in the abstract and introduction implies that the study is specific to this application, which is not.

2- Your model was fed with data of deceased patients (100% in 5 years), but the human agents (doctors) were not sure about death of patients within the next 5 years period (92% for the next 6 month in hospice). I wonder if this difference should be considered when comparing the results (it might slightly favor the machine learning approach) ?

3- The median survival for your reference of correct prediction (ref [22]) was 24 days. As the criterion of correct prediction (33%) gets more generous for longer survivals, then I think comparing accuracy of predictions to the reference is acceptable only if the median survivals are comparable . I could not find median survival days in the manuscript, please indicate that, and if it is very different from your reference you may want to clarify why still you compare the results while the survival median in your study is longer (if longer).
 Minor:

1- P7L9-13: please provide reference for (probably #22 ?)

2- P7L20: please discuss why the results should be specific to ACP or if it can be generalized

3- P16L44: do you mean "forward stepwise " feature selection approach? As you did not tested all combinations, then you may want to motivate why you preffered this approach.

4- P27L9-13: this is an important extension to your work, you may want to elaborate more on the requirement of this "more data"

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 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.

No

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