Reviewer's report

Title: Applying Machine Learning to Predict Future Adherence to Physical Activity Programs

Version: 0 Date: 03 Jul 2019

Reviewer: Slade Matthews

Reviewer's report:

I commend the authors on a very nicely prepared piece of work. The methodology was sound and clearly articulated. The results were very promising indeed and the trends that were uncovered in terms of the predictive power of features included in the LR model made intuitive sense. The use of the data from the first 15 weeks to predict subsequent activity made for a nice separation of training and test data preventing any bleed across beyond that which would be brought about by a repeated measures design if you like.

I have only a very small number of improvements in the communication in the paper viz:

Page 4, line 8 change "Americans" to American

Page 6 line 2 change "conduct" to conducted

Page 16 line 13 - Here in the Discussion you talk about the predictive features that were significant in the LR model. "The most predictive features in DiPS are: week number, steps data, inculding... -- Here I think you should put the steps data in parentheses rather than putting a comma after data so that the terms "steps data" is clearly referring to the subsequent features. then the and physical activity data will be outside parentheses and clearly refer to the other kinds of features that were significant.

There were my only minor changes that I would like to see. Overall I think you have done a great job communicating this study and the results are very promising. I did ask the question in my head while reading the paper "Why not other algorithms such as decision trees or random forests?" But I noted that you did address this question in the results section and I suppose that the results were so good with LR and SVM that it probably didn't seem necessary to present other algorithms. Nonetheless, I would still be interested in the performance of algorithms that take a fundamentally different approach to modelling the answer space.
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.