Reviewer Report

Title: Deep learning for clustering of multivariate clinical patient trajectories with missing values

Version: Original Submission **Date:** 8/14/2019

Reviewer name: Karsten Borgwardt

Reviewer Comments to Author:

The work introduces a clustering method for multivariate time series with missing values using an autoencoder architecture. The method is successfully applied to stratify patients according to their trajectories, which is an important problem.
This is a very good article, which I enjoyed reading. The write-up is well structured, the main message is clear and the results are relevant. I particularly enjoyed that all experimental results are reported with an error assessment and that a simulation experiment is conducted.
I have one suggestion to further improve the simulation experiment, which is increasing the number of comparison partners: 1) Comparing VADER to VaDE with an explicit and/or no imputation scheme. This would quantitatively assess the bias of separating clustering from imputation. 2) Compare hierarchical clustering with a multidimensional DTW distance measure. DTW is a classic and successful distance measure in the data mining community.

The paper is well written. To my understanding, aside from DTW, all relevant related work has been cited.

Level of Interest

Please indicate how interesting you found the manuscript: Choose an item.

Quality of Written English

Please indicate the quality of language in the manuscript: Choose an item.

Declaration of Competing Interests

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

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

- 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?
- Do you have any other financial competing interests?
- 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.

Choose an item.

To further support our reviewers, we have joined with Publons, where you can gain additional credit to further highlight your hard work (see: https://publons.com/journal/530/gigascience). On publication of this paper, your review will be automatically added to Publons, you can then choose whether or not to claim your Publons credit. I understand this statement.

Yes Choose an item.