

Review of: "Implementing Machine Learning to predict the 10-year risk of Cardiovascular Disease"

Preesha Premsagar¹

¹ University of KwaZulu-Natal

Potential competing interests: No potential competing interests to declare.

Dear Editors

Thank you for the opportunity to review this relevant and important piece of research.

These are my comments below:

1. Overall this is a very well written paper, because it is well structured, follows logical methodology and addresses a research question in a clear and precise manner.
2. This is relevant research because it is current and has clinical significance worldwide, and could potentially have massive implications in decision making.

My recommendations are

1. A few simply explanations on the input data may help to improve the understanding of non-medical people. In the case of the types of chest pain, the authors have explained well. However, for example, a few sentences to explain thal (which is a thallium scan, that is a type of nuclear scan), as well as using full words where possible (chol for cholesterol) can help ther reader identify the what the data represents.
2. In the heat map, for the correlation matrix, if possible, please write out the variable names in full. This need not be re-running the codes, this can be done by editing the labels in Paint or another software, that keeps the results 100% intact and just allows the author to write the labels in full. This also helps the readers to know what variable correlated well with each other.
3. If possible, please add numbers to the bars of the histograms in figures 3, 5 and 6. PowerPoint is good for this.
4. Please discuss class imbalance and fitting in the introduction and discussion.
5. Please use a few sentences to explain ensemble method as some readers may not be familiar with this.
6. Lastly, please provide some very basic statistics of morbidity and mortality rates of cardiovascular disease. This research has 2 very strong components - one is that machine learning is current and cutting edge, and the other is that the disease area of cardiovascular disease is a leading cause of morbidity and mortality. It is important to tie the clinical significance of this disease area into the paper.

I think that this paper should be considered for publication after changes are made. Congratulations to the authors on good work.

Kind regards,

Reviewer