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

Saman Rajebi

Potential competing interests: No potential competing interests to declare.

- In section 3.1, in the data description, the number of classes and the number of members of each should be specified.
- In section 3.2, if the division of data into training and testing is not done randomly, the accuracy of the results will be doubtful. If this division is done randomly, it should be mentioned.
- The title of Figure 3 should be changed. For example, the term "feature ranking" can be used.
- In the feature selection, only the correlated features have been removed. Feature combination methods such as FDR and PCA could have yielded better results.
- For the methods used, additional explanations are not provided. For example, in the case of the KNN method, the results based on which k value was calculated and was this value the optimal value or not? Or for the SVM method, has optimization been done on the SVM parameters, such as the box constraint.
- The selected parameters on the random forest method and its comparison with the decision tree results can have valuable results in data analysis.
- In new classifications, it is more focused on the combination of classifications with each other. Instead of choosing one classifier, it is better to work on the simultaneous use of several classifiers.