Review of: "Chronic disease treatment default prediction with random sampling optimization."

Seyed Ehsan Roshan[1]

1 University of Tehran

Potential competing interests: No potential competing interests to declare.

This paper presents an experimental assessment of methods using healthcare datasets. My comments are as follows:

1. In the abstract, it is necessary to clarify the main goal and the advantages of your work.

2. The introduction section requires revision and rewriting. Please refer to your contributions, your goals, and express your research question clearly.

3. It would be beneficial to explain the limitations of the methods in section 1.1.

4. Why did you choose to utilize the following classifiers? What was the rationale behind your decision to use "Gradient boosting, Extreme gradient boosting, Logistic regression, Support vector machines, and Random forest"?

5. Section 2.2 needs to be explained more clearly. How do you handle null values? How do you deal with outliers? Which normalization techniques are applied? How are hyperparameters adjusted?

6. In the results section, I recommend using precision-recall AUC. Additionally, you can create synthetic imbalanced datasets with a high imbalance rate to analyze the behavior of the methods.