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

Roselina Sallehuddin[1]

1 Universiti Teknologi Malaysia

Provide some discussions on the issues encountered while handling the imbalance dataset: for example, the effect of the imbalance dataset on the accuracy performance, limitations of the existing random sampling methods, and the need to do the assessment.

List down your research questions and the research objectives that reflect each other.

Revise the format for Figure and Table labeling. In Figure 2.1 there are only 2 classes for the output, but in Figure 2.3, there are more than 2 classes, please explain.

Figure 2.3 demonstrates the research methodology for this study. Which part is related to random sampling? Which one is under or over-sampling? In the figure, there are, outliers, dimension reduction, feature selection, and parameter tuning which are not been discussed in the previous part. Can you explain, how these activities are related to your research questions and objectives?

What are the purposes of Figure 2.4 and Figure 2.5?

The author needs to revise the flow and the content of the paper. The main topic here is how to investigate the impact and effect of random sampling use in predictive modeling. But which part of the paper explains and demonstrates how investigation, comparison, and evaluation are made?

In the title, there is optimization, what are going to optimize here? It is not clearly discussed and mentioned in the paper.