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

Toshitaka Hayashi[1]

1 University of Hradec Králové

The authors should address two major concerns as follows:

1. In Figure 2.3, when is the sampling stage?

-It is important to note that sampling must be done AFTER the train-test split.

-Testing data must not include any synthetic samples.

-The authors should re-run the experiments if they did sampling before the train-test split.

2. Figure 2.4 includes several issues.

-There are unclear arrows. What is the meaning of the arrow from random forest to K-NN?What are the arrows from decision tree, naive bayes to SVM?

-The same comment is given for regression algorithms.

-Logistic regression is a classification technique.