Review of: "Bank Customer Churn Prediction Using SMOTE: A Comparative Analysis"

Majid Danesh[1]

1 Sari Agricultural Sciences and Natural Resources University

Potential competing interests: No potential competing interests to declare.

Dear Author(s),

At the outset, I appreciate your valuable research. As far as I know, the modeling methods, and it is part of my expertise, it would be better to consider these tips to enrich and scientifically improve the quality of your manuscript. Hence, I draw your attention to the following points:

- Please explain thoroughly in the introduction how ML approaches can handle unbalanced data (as the strength of ML algorithms and data mining methods).

- It would be great if you could highlight and emphasize the novelty of the research in the abstract, introduction (after objectives), and final discussion sections (the essentialness of doing the research).

- Is there any separation/division of data into calibration (train)/validation (test) subsets? If so, please specify them correctly (in the materials & methods).

- In practice, does the SMOTE algorithm only apply to manage and control the data imbalance, or more than that? Please mention in the relevant section.

- In which environment (software) did you run the algorithms and build the models (SMOTE, GA, KNN, RF, AdaBoost, ANN)? (Please include its name(s) in the relevant section.)

- What was the exact number of samples? Also, the number of samples in each split subset (train & test datasets).

- What metrics were precisely used for ROC-AUC analyses? In which environment are these metrics extracted (software name)? For better comprehension of the readers, further interpretation in the pertinent section is appreciated.

I wish you all the best and look forward to your next worthy research.

Majid Danesh

Assistant Prof., Dept. of Soil Science,

SANRU

IRAN