Review of: "Predicting Mobile Money Transaction Fraud using Machine Learning Algorithms"

Albert Buabeng

Potential competing interests: No potential competing interests to declare.

Dear Authors,

In the manuscript, Machine Learning (ML) algorithms have been implemented to predict/classify mobile money transactions whether they are fraudulent or not. The authors generated synthetic data for validating their proposed techniques. The manuscript is understandable and presents some experimental results. However, in order to meet the scientific standards, some comments should be addressed to revise the manuscript.

- The text of the manuscript in general needs a thorough review as there are multiple referencing errors, especially for Tables, Figures and Equations. This affects the clarity of information being disseminated. Also, some Tables and Figures in the manuscript have not been referenced in the manuscript. Again, to improve readability, the authors should consider discussing the result followed by the table or figure discussed. At this state, readers will have to be flipping through the pages.
- The authors presented a shallow review of literature in Section 2. The review of literature should also provide information on what other research works have been conducted on the subject, their findings, the gaps identified and possible the motivation for the current work. Yet the authors failed to do so.
- During the data processing stage, the authors adopted SMOTe-ENN as a resampling strategy. Though this is the right approach, the sequence should be considered. Normally, the SMOTe-ENN is implemented on the training data set, not the entire 1048575 rows of data synthetically generated. This is because the test data should be in its original state so that the credibility and generalizability of the technique can be ascertained.
- Under the results section, the experimental results shown need to be reorganized in a compact form. In this current state, they have been split into several tables which in my opinion could be put together. This also portrays that the authors do not fully grasp the concept presented as the experimental results.
- The ML algorithms utilized requires in the manuscript require the initialization of some parameters in order to optimally train the dataset. The authors should come clear on how these were achieved.
- The authors later in the results section indicate that the data was split into training and testing, what was the ratio?
- Are there other limitations of your study? Clarifying the limitations of a study allows the readers to understand better under which conditions the results should be interpreted. A clear description of the limitations of a study also shows that the authors have a holistic understanding of their study.
- The discussion of the results needs to include the strengths and weaknesses of the proposed techniques.
- The conclusion should be crisp and concise. The authors should discuss the pros and cons of the study in the Conclusion section.