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

Abida Toumi[1]

1 Université de Biskra

Potential competing interests: No potential competing interests to declare.

Grammar and Wording:

The grammar and wording of the article are generally clear and concise. There are a few minor grammatical errors and awkward phrasings, but they do not significantly impact the understanding of the content. For example,

- In the abstract, the phrase "some machine learning (ML) algorithms have struggled to deliver the required performance in identifying customer churn accurately, most especially when the dataset is imbalance data" is unclear and could be improved.
- In the abstract, it would be better to say "evaluated using" instead of "evaluate using." Overall, the language used is appropriate for an academic journal.
- The title of section 3.3, ‘Genetic’ algorithm, should be 'Generic' Algorithm.

Logic and Science:

- The article provides a logical and well-structured explanation of the problem of customer churn in the banking industry and the use of machine learning techniques to predict churn.
- The introduction provides a clear background and rationale for the study. The authors accurately identify the challenges of class imbalance and high-dimensional data in churn prediction and propose solutions, such as the use of SMOTE and genetic algorithms.
- However, there are some gaps in the logic and explanation of the research. For example, the authors briefly mention the use of Synthetic Minority Over Sampling Technique (SMOTE) and Genetic Algorithm (GA) but do not provide sufficient details on how these techniques were applied. This lack of clarity can influence the evaluation, the validity, and the reliability of the study.
- The related works section provides relevant studies and their findings, demonstrating a good understanding of the existing literature.

Innovation:

The article contributes to the field of customer churn prediction in the banking industry by proposing the use of SMOTE and genetic algorithms to address class imbalance and high-dimensional data issues. This approach is innovative and has the potential to improve the accuracy of churn prediction models. The authors also compare their results with existing

models and demonstrate the superior performance of their proposed strategy.

Weaknesses:

- The article lacks a clear research objective or hypothesis. It would be beneficial to explicitly state the research question the study aims to answer.

- The methodology section is not comprehensive. It would be helpful to provide more details on the dataset used, the specific SMOTE and genetic algorithm techniques employed, and the evaluation metrics used for comparing the classification algorithms.

- The limitations of the study are not adequately discussed. It would be valuable to address potential limitations, such as the generalizability of the findings to other banking institutions and the impact of data imbalances on classifier accuracy.

General Evaluation:

Overall, the article provides a clear overview of the problem of customer churn and proposes an innovative approach using SMOTE and genetic algorithms for prediction. However, the article could be strengthened by providing more details on the dataset, justifying the choice of techniques, and presenting a more comprehensive analysis of the results. Additionally, the logic and organization of the introduction, methodology, and results sections could be improved for better clarity and coherence. If all remarks are considered, the paper can be accepted.