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

Farah Jemili[1]

1 Université du Centre, Sousse

-While the use of Synthetic Minority Over Sampling Technique (SMOTE) is a good approach to handle imbalanced data, the authors should also consider other techniques such as undersampling the majority class or using cost-sensitive learning. For example, they could use different weights for the classes in the loss function to make the model pay more attention to the minority class.

-The use of Genetic Algorithm for feature selection is interesting, but the authors should provide more details about the parameters used in the GA, such as population size, mutation rate, and crossover rate. They should also discuss why they chose GA over other feature selection methods.

- The authors should provide more details about the hyperparameters used in the K-Nearest Neighbor (KNN) model, which achieved the best performance. They should also discuss the possibility of overfitting, given that KNN can be sensitive to the choice of the number of neighbors and the type of distance metric used.

- The authors should provide a more detailed comparison with existing studies. For example, they could include a table that summarizes the main characteristics and results of the compared studies. They should also discuss the reasons why their model outperformed the others.

-While SMOTE is a popular method for dealing with class imbalance, the authors should discuss its limitations and potential issues, such as the risk of over-generalization and the introduction of noise.

-The authors should provide more information about the dataset used, such as its size, the number of features, and the distribution of the classes. They should also discuss the limitations of the dataset and how they might affect the results.

-The authors should justify the choice of evaluation metrics used in the study. They should also consider using additional metrics that are more suitable for imbalanced datasets, such as the Area Under the Receiver Operating Characteristic Curve (AUC-ROC).