5. CONCLUSION

5.1 Conclusion

This research focuses on comparing the effectiveness of six algorithms, which are divided into two types: single classifiers and ensemble learning. The three types of single classifiers are Decision Trees, Naive Bayes, and Logistic Regression. Whereas, ensemble learning consists of AdaBoost, Random Forest, and XGBoost. Throughout this study, the focus also examines the impact of having the imbalanced data and the exploration of how Synthetic Minority Oversampling Technique(SMOTE)-NC tackled the imbalance problem.

The comparison and assessment of six algorithms is based on performance indicators, specifically F1-Score and G-Mean. The decision to consider these two performance indicators is due to the issue of imbalanced data. The F1-Score is a crucial metric that measures the balance between precision and recall, providing insight into how well the model performs in identifying positive cases while minimizing false positives. The G-Mean also measures the balance between specificity and sensitivity, which is crucial for addressing imbalance problems. The final comparison results indicate that Random Forest, an ensemble learning method, exhibits optimal performance compared to the other five algorithms, with an F1-Score of 53.7%, and a G-Mean of 69.1%. The AUC for this method reaches 77.3%, which can be considered as a good performance.

The challenge of imbalanced data is particularly prominent when assessing the risk of credit card default, as the minority class (default customers) is underrepresented. Therefore, this research also observes the impact of having imbalance data in classifying the credit card default payments and the synthetic minority oversampling technique for numerical and categorical data, known as SMOTE-NC, has proven to be an effective method for improving model performance across various imbalance ratios. In cases where the imbalance ratio is greater, SMOTE-NC can have a greater impact on performance. Furthermore, SMOTE-NC helps the model prevent prediction failures, especially in scenarios with extremely high imbalance ratios, when the absence of oversampling techniques could result in inaccurate and failed predictions.

Another highlight of this research is the features importance in predicting the credit card default to minimize the risk. The factors such as the customer's repayment status in the last one until four months along with the credit card’s limit need to be considered more.

Practical recommendations are provided to financial institutions to carefully consider these variables in developing a more effective and accurate credit risk prediction system. In addition, the decision to use the SMOTE or apply the data directly to the model depends on the dataset characteristics. If the dataset is highly imbalanced, applying the SMOTE is necessary, but if the imbalance ratio is not significant, oversampling can be skipped. When the dataset is large and the model can understand its pattern, applying SMOTE is also unnecessary. However, if the dataset size is small, SMOTE is needed to help generalize the model. In order to manage risk and sustain the credit card customer portfolio, SMOTE is relevant for financial institutions' goals and policies.

5.2 Recommendation

The recommendation for further research is to explore finding the best hyperparameters using different optimization techniques, allowing the discovery of better hyperparameter combinations to enhance the model's performance in future predictions. Additionally, experimenting with the development of a more robust model by combining various algorithms could be valuable for further investigation. It is also suggested to assess the practical implementation of the developed model by expanding the analysis using detailed performance metrics. This approach will help in understanding the overall capability and performance of the model in addressing real-world challenges.

In order to enhance the model’s performance for better generalization and improve the prediction, it might be needed to consider the addition of other factors that haven’t been represented in this dataset. The presence of these new factors can potentially contribute to achieving a better harmony in the model, especially in terms of F1-Score. New features that could have a substantial impact such as monthly income, providing valuable insights into the customer's repayment capability, and the number of monthly transactions. Analyzing the frequency with which customers utilize their credit cards each month can offer the card issuing institution a deeper understanding of customer behavior. By considering these factors, the model can be fine-tuned to better adapt to real-world scenarios, ultimately leading to more accurate predictions and a more robust evaluation of credit card users' financial behavior.

Despite the machine learning's capability in predicting credit card defaults, research and development in machine learning can be expanded to assist card issuing institutions in other areas, such as the frequent occurrence of credit card fraud detection. It all comes back to the concept of machine learning where the model learns from existing data and enhances its ability to generalize previously unrecognized data. The trained model is expected to make accurate and precise predictions, thereby supporting the performance of financial institutions.