

Module 12 Report Template

Overview of the Analysis

In this section, describe the analysis you completed for the machine learning models used in this Challenge. This might include:

- * Explain the purpose of the analysis.
- * Explain what financial information the data was on, and what you needed to predict.
- * Provide basic information about the variables you were trying to predict (e.g., `value_counts`).
- * Describe the stages of the machine learning process you went through as part of this analysis.
- * Briefly touch on any methods you used (e.g., `LogisticRegression`, or any other algorithms).

Results

- Machine Learning Model: Logistic Regression
 - Training Results:
 - Accuracy: 99%
 - Precision for Healthy Loans (0): 1.00
 - Recall for Healthy Loans (0): 0.99
 - Precision for High-Risk Loans (1): 0.85
 - Recall for High-Risk Loans (1): 0.98
 - AUC: 0.9942
 - Testing Results:
 - Accuracy: 99%
 - Precision for Healthy Loans (0): 1.00
 - Recall for Healthy Loans (0): 1.00

- Precision for High-Risk Loans (1): 0.88
 - Recall for High-Risk Loans (1): 0.98
 - AUC: 0.9970
- Machine Learning Model: XGBoost Classifier
 - Training Results:
 - Accuracy: 99%
 - Precision for Healthy Loans (0): 1.00
 - Recall for Healthy Loans (0): 0.99
 - Precision for High-Risk Loans (1): 0.85
 - Recall for High-Risk Loans (1): 0.99
 - AUC: 0.9978
 - Testing Results:
 - Accuracy: 100%
 - Precision for Healthy Loans (0): 1.00
 - Recall for Healthy Loans (0): 1.00
 - Precision for High-Risk Loans (1): 0.88
 - Recall for High-Risk Loans (1): 0.99
 - AUC: 0.9977
- Machine Learning Model: K-Nearest Neighbors (KNN)
 - Training Results:
 - Accuracy: 99%
 - Precision for Healthy Loans (0): 1.00
 - Recall for Healthy Loans (0): 0.99
 - Precision for High-Risk Loans (1): 0.85
 - Recall for High-Risk Loans (1): 0.99
 - AUC: 0.9970

- Testing Results:
 - Accuracy: 100%
 - Precision for Healthy Loans (0): 1.00
 - Recall for Healthy Loans (0): 1.00
 - Precision for High-Risk Loans (1): 0.88
 - Recall for High-Risk Loans (1): 0.99
 - AUC: 0.9971
- Machine Learning Model: LightGBM (LGBM)
 - Training Results:
 - Accuracy: 99%
 - Precision for Healthy Loans (0): 1.00
 - Recall for Healthy Loans (0): 0.99
 - Precision for High-Risk Loans (1): 0.85
 - Recall for High-Risk Loans (1): 0.99
 - AUC: 0.9982
 - Testing Results:
 - Accuracy: 100%
 - Precision for Healthy Loans (0): 1.00
 - Recall for Healthy Loans (0): 1.00
 - Precision for High-Risk Loans (1): 0.88
 - Recall for High-Risk Loans (1): 0.99
 - AUC: 0.9975

Summary

The Four models were evaluated: Logistic Regression, XGBoost, K-Nearest Neighbors (KNN), and LightGBM (LGBM). All four models performed well, with accuracy between 99–100% and AUC scores above 0.99, to provide a strong ability to distinguish between healthy and high-risk loans. The model maintains a rate of 98% on both training and test data. Precision is slightly lower 0.85

on training, 0.88 on testing which indicates the small number of healthy loans are predicted as high-risk. The performance is consistent across test sets, the model does not overfit and generalizes well. The models like XGBoost and LGBM, identified risky loans while maintaining great performance on healthy loans.