Credit Risk Analysis Report

Overview of the Analysis

This analysis is to evaluate a machine learning model for credit risk classification to build a model that can identify the creditworthiness of borrowers. The dataset contains financial information about loan applicants, and our goal is to forecast whether a loan is high or low risk based on given data. A machine learning analysis will identify the possible risk of loans which is good or risky loans.

Process:

1. Data Preparation:

- The dataset was loaded and preprocessed.
- The target variable (loan status) was separated from the features.

2. Model Selection:

- Logistic Regression model for classification.

3. Model Evaluation:

- We examined performance using accuracy, precision, and recall metrics.

Results

Machine Learning Model: Logistic Regression

- Accuracy: >99%
- Precision (for class 1): 84%
- Recall (for class 1): 94%
- F-score (for class 1): 89%
- Precision 100%, Recall 99%, and F-Score are all 100% for class 0.

Summary

The accuracy level was more than 99% so we can say that the logistic regression model performed very well. Given the nature of credit risk classification, recall is particularly important as it ensures we correctly identify as many high-risk loans as possible. With a recall score of 95%, the model effectively captures most of the high-risk cases, making it a viable choice for credit risk assessment.

Recommendation:

- As Given its high accuracy and strong recall performance, we recommend this model to use for credit risk classification.