Credit Risk Analysis Report

Overview

This analysis evaluates a machine learning model for credit risk classification to assess borrower creditworthiness. The dataset contains financial details of loan applicants, and the goal is to predict whether a loan is **high-risk or low-risk** based on the provided data. Machine learning helps identify potential risks, ensuring informed lending decisions.

Process

1. **Data Preparation:**
    - Loaded and pre-processed the dataset.
    - Separated the target variable (loan status) from the features.
2. **Model Selection:**
    - Chose **Logistic Regression** for classification.
3. **Model Evaluation:**
    - Assessed performance using **accuracy, precision, recall, and F1-score**.

Results

Machine Learning Model: Logistic Regression

- **Accuracy:** >99%
- **Precision (High-Risk Loans):** 84%
- **Recall (High-Risk Loans):** 94%
- **F1-Score (High-Risk Loans):** 89%
- **Precision (Low-Risk Loans):** 100%
- **Recall (Low-Risk Loans):** 99%
- **F1-Score (Low-Risk Loans):** 100%

Summary

With an accuracy of over **99%**, the **logistic regression model performs exceptionally well** in credit risk classification. Given the critical nature of identifying high-risk loans, **recall is a key metric**, ensuring that most risky loans are correctly classified. With a **94% recall for high-risk loans**, the model effectively captures the majority of potential defaulters, making it a **reliable tool for credit risk assessment**.

Recommendation

Given its **high accuracy and strong recall performance**, we recommend using this model for **credit risk classification**, as it provides a balanced approach to identifying both **healthy and high-risk loans**.