Purpose of the analysis

The purpose of the analysis is to apply machine learning models to predict credit risk, to identify if an applicant would be either low risk or high risk for credit loans. This approach would help the companies to make better decisions on loan approvals.

The data provided information related to loans applicants and the target variable to predict is “loan_status”

- 0 (Healthy Loan): The applicant is likely to repay the loan without defaulting.
- 1 (High-Risk Loan): The applicant is at high risk of defaulting on the loan.

The process to obtain the results were

- Separated the data into labels and features
- Split the data
- Train the model using logistic regression model
- Evaluate and analyze the results

Results

- Accuracy Score: 99%
- Precision Score for Class 0 (Healthy Loans): 1.00
- Recall Score for Class 0 (Healthy Loans): 1.00
- Precision Score for Class 1 (High-Risk Loans): 0.86
- Recall Score for Class 1 (High-Risk Loans): 0

Recommendation

Based on the results from the analysis, I would recommend using a logistic regression model to determine whether an applicant is a healthy loan or high risk since the model provided 99% accuracy which makes it reliable, this would help minimize the risk of financial losses from the companies