

Overview

In this challenge I was tasked to train and evaluate a model to predict loan risk. By using the logic regression model, I was able to display my results and determine if they were “healthy” or “unhealthy”. The dataframe that I read in from the csv provided had other relevant information such as loan size, interest rate, borrower income, debt to income ratio, number of accounts, derogatory remarks, and total debt. Initially there was our target (y) listed as a column. I was able to drop that column in order to separate it from the (X) variable. From there I split the data into X and y train and X and y test. I made sure to use the random state of 1 in order to be able to reproduce the work. I trained the data using the Logistic Regression model. From there I generated a confusion matrix then displayed the classification report for the model.

Results

- The accuracy of this model is 99%. This is a fantastic result as it was 99% successful in its classification.
- The “healthy” loans precision is at 100%
- The “unhealthy” loans precision is at 85% which is still great seeing as the support data is only 619 where as the healthy data was given 18765.
- The “healthy” recall is at 99%
- The “unhealthy” recall is 91%

Summary

This logistic regression model was very successful and I would trust it. The Accuracy , precision and recall numbers are very high which shows this being a dependable model.