This assignment aimed to create a machine-learning model that can be used to predict if an applicant would result in a high-risk or healthy loan. After reviewing the data provided regarding the loans, including interest rates, amounts, and information about the borrower’s debt and accounts, it was determined that only 2,500 of the total loans were classified as high-risk.

The results using the model are as follows:

Accuracy: 99%

Balanced Accuracy Score: 95.2%

Precision:

* Healthy: 1.00
* High Risk: 0.85

Recall:

* Healthy: 0.99
* High Risk: 0.91

While the Logistic Regression model shows a high rate of accuracy, I cannot be certain that this is the perfect model to determine if a model would be a healthy or high-risk loan. The model shows that approximately 10% of high-risk loans are reported to be healthy, leaving a lender at risk for approving loans that would quickly end in default. Further, there were not as many high-risk loans in the data set for this model to learn.