Credit Risk Analysis Report :

**Provide an overview that explains the purpose of this analysis. (5 points):**

The purpose of this analysis is to evaluate the performance of the machine learning model. The goal is to see if it has good predicting capacities. We hope that the model can create a good output based on what we inputed. This is why its supervised learning. First we separate the y and x variable so that we can separate the data into 2 data sets so that we can compare the output , by this we also get to have better results because we test it on 2 different sets so it can be more accurate.it randomly split the data into a testing and training data. We only use the training data to then test on the test data and then compare their result. Then we fit a logistic regression model so that we can train the machine to understand the correlation between the 2 variables. Then we do a confusion matrix to see how the model performs on a classification task and see what type of error it generates. From there, we get the classification report that allows us to predict the accuracy of the model capabilities to predict. The closer it is to 1 the better it is at predicting outcomes.

**Using a bulleted list, describe the accuracy, precision, and recall scores of the machine learning model. (5 points)**

* The accuracy is the ratio that predics how good the predictions of the model is.The higher the accuracy the better the model.
* The precision is = positive predictions/total positive predictions, if it high it means that its very accurate
* If the Recall is high then it means that the model can accurately predict most postives .it measures what percentage of the positive were we able to detect.

**Summarize the results from the machine learning model. Include your justification for recommending the model for use by the company. If you don’t recommend the model, justify your reasoning. (10 points)**

When we made the prediction, the actual and prediction were accurate and gave the same results. 1-1

precision recall f1-score support

0 1.00 1.00 1.00 18759

1 0.87 0.90 0.89 625

accuracy 0.99 19384

macro avg 0.94 0.95 0.94 19384

weighted avg 0.99 0.99 0.99 19384

18759 is the number of rows of data set that were 0 and 625 is the number of data sets that were 1. It’s the distribution of 1 and 0s.

The accuracy is 0.99 on 19384 data sets and the precision is 0.87.

confusion\_matrix(y\_test, predictions)

array([[18678, 81],

[ 62, 563]])

Based on the confusion matrix we can conclude that we had 18678 True negative which means that it correctly predicted 18678 of its data sets.

81 False positive

62 False negative

563 True postive

This combined with an accuracy of 0.99 and a precision of 0.89, I would recommend the model. Even the F1 is high with 0.89 which means that the model has a good balance between precision and recall.