**Module 20-Supervised-Learning Report**

**Overview of the Analysis**

The purpose of this analysis was to try to predict if a borrower would default on a loan or not based on factors such as: loan size, interest rate, borrower income, debt.

The goal was to predict loan status.

The data, which was in a csv, was read into a Pandas data frame, then split the data into training and testing data groups, and then fit to various models and tested their prediction accuracy.

Models used were logistic regression and random forest. Then the data was scaled and fit to a logistic regression model again.

**Results**

Using bulleted lists, describe the balanced accuracy scores and the precision and recall scores of all machine learning models.

- Machine Learning Model 1
    - Logistic Regression: Accuracy - 99%, Precision - 87%, Recall - 89%
- Machine Learning Model 2
    - Knn: Accuracy - 100%, Precision - 100%, Recall - 100%
- Machine Learning Model 3
    - Random Forest: Accuracy - 100%, Precision - 100%, Recall - 100%

**Summary**

The model I recommend is a logistic regression on scaled data. Both the Knn and Random Forest models seemed to be overfitting with perfect scores which is suspicious.