Credit Risk Analysis

Purpose:

The purpose of this analysis is to predict credit risk. We are using a credit card dataset from LendingClub, a peer-to-peer lending service to complete our analysis. We are tasked with using several different statistical models to evaluate the data and try to figure out which model performs the best at predicting credit risk.

* Results:

**Balanced accuracy scores:**

Balanced Accuracy is the mean of Sensitivity and Specificity. Where Sensitivity (True Positive Rate) is the probability of a positive case being accurately classed as being positive, and Specificity (True Negative Rate) is the probability of a negative case being accuracy classed as negative.

 Rule for 'good' scores is:

Over 0.9 - Very good

Between 0.7 and 0.9 - Good

Between 0.6 and 0.7 - OK

Below 0.6 - Poor

1 RandomOverSampler : 0.6293939430565123

2 SMOTE : 0.6277008271188627

3 RandomUnderSampler : 0.5902034962189427

4 SMOTEENN : 0.6411460410698961

5 Random Forrest : 0.9945945945945946

6 GradientBoostingClassifier: 0.9945945945945946

**Precision scores:**

1 RandomOverSampler : 0.6293939430565123

2 SMOTE : 0.6277008271188627

3 RandomUnderSampler : 0.5902034962189427

4 SMOTEENN : 0.6411460410698961

5 Random Forrest : 0.9945945945945946

6 GradientBoostingClassifier: 0.9945945945945946

**Recall scores:**

1 RandomOverSampler : 0.6293939430565123

2 SMOTE : 0.6277008271188627

3 RandomUnderSampler : 0.5902034962189427

4 SMOTEENN : 0.6411460410698961

5 Random Forrest : 0.9945945945945946

6 GradientBoostingClassifier: 0.9945945945945946

* Summary:
  + There is a summary of the results (2 pt)
  + There is a recommendation on which model to use, or there is no recommendation with a justification (3 pt)