Overview

For this analysis, the purpose is to use supervised machine learning to evaluate credit risk. I evaluated the data using scikit-learn and imbalanced-learn libraries.

Results

Ensemble Learning with BalancedRandonForestClassifier

* Accuracy 78%
* Precision 99%
* Recall 87%

Ensemble Learning with EasyEnsembleClassifier

* Accuracy 91%
* Precision 99%
* Recall 90%

Combination Sampling with SMOTEEN

* Accuracy 67%
* Precision 99%
* Recall 66%

Undersampling with ClusterCentroids

* Accuracy 67%
* Precision 99%
* Recall 66%

Oversampling with SMOTE

* Accuracy 67%
* Precision 99%
* Recall 66%

Naïve Random Oversampling with RandomOverSampler

* Accuracy 67%
* Precision 99%
* Recall 66%