MISM6203 Project 3

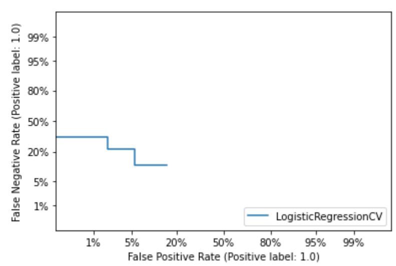
Michael Conan

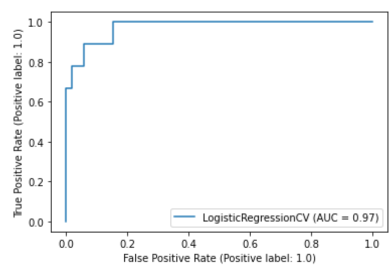
Credit Card Fraud Detection

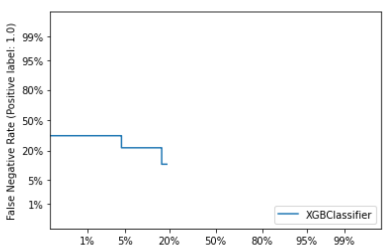
In order to effectively analyze NextLEVEL Bank Corp's credit card transactions for fraudulent activity, I developed 2 binary classification models. These models are trained using historical transaction data to statistically compute the probability of a transaction being fraudulent. To improve the performance of the alogrithms, I tested changing the distribution of fraudulent transactions (oversampling) and incorporating cross-validation to train the multiple multiple times on different portions of the data.

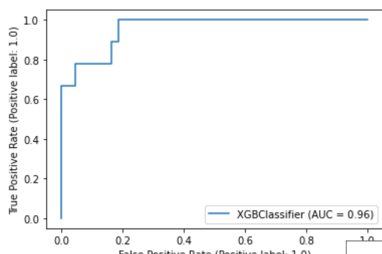
For each the Logistic Regression and XGBoost algorithms, I tested 4 different scenarios. The first was using the original distribution and the full training dataset. The second incorporated 5-fold cross-validation with the original distribution. The third scenario included the oversampled dataset with the full training data, and the final included cross-validation and oversampling. For the XGBoost, I additionally tested multiple sets of parameters for the maximum tree depth and alpha regularization parameter. I chose to evaluate these models using the F1 score, which is a weighted average metric of precision and recall, which both focus on the ability of the model to effectively predict fraudulent entries.

For the Logistic Regression model, the first scenario was the most effective and achieved an F1 score of 0.71 for the testing set. The accuracy was quite high with an AUC in the receiver operating curve of 0.97.

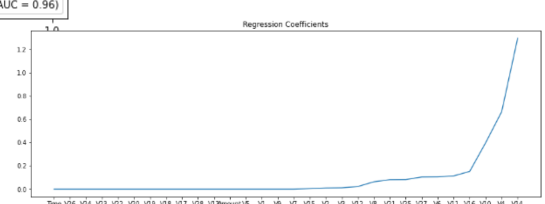




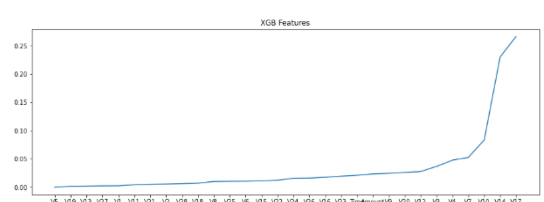




The XGBoost performed even better with an F1 score of 0.8. The first scenario was again the best performing of the 4. The AUC of the ROC was actually lower with a value of 0.96 compared to the Logistic Regression model trained.



In comparing the features in the two models, the logistic regression assigns more weight to less important variables, while the XGBoost assigns the majority of the prediction to the top few predictors. The most important predictors are slightly different between the models, but both include V10 and V14 as top factors. No context was provided on what these values represent, but they are likely optimized features from the underlying transaction attributes.



In calculating the economic costs of the model assuming $4,000 per type II error (fraud missed) and $25 per type I error (false alarm), the XGBoost resulted in equal or lower costs by up to 33% based on the threshold applied. This model is able to effectively balance the costs based on the use of F1 score to effectively identify the fraudulent entries in the population.