

CALLMINER

Predict Debt Collection Default Risk

THE CHALLENGE

A US collection agency had a very high rate of debt collection defaults. A high default rate is rather typical in the collections industry.

In 2017, the Consumer Financial Protection Bureau (CFPB) estimated that 70 million Americans had, or were contacted about, a debt in collection during the prior year. Out of 42 million student loans, one in four are delinquent.

In 2016, 18% of Americans with credit reports had medical debt in collection. Collection agencies make more than a billion collection calls a year, sometimes dozens of calls per account per month. 15% of debtors contacted by collection agencies are sued every year.

Naturally, collection agencies are seeking for ways to reduce their collection effort failures, as was the collection agency requesting this evaluation.

The agency knew that consumers generally default on debt due to unemployment, illness, divorce, or other unanticipated hardships.

However, they assumed that the defaults are also related to the behavioral characteristics of the borrowers, such as higher risk-taking, impulsive behavior, low personal integrity, bad repayment habits, lack of planning, etc.

They wondered whether assessing personal tendencies at the beginning of the collection process could predict the collection outcome. Predicting potential high and low risk debtors could help focus their collection efforts and help adjust their approach to each debtor.

THE SOLUTION

The agency's collection process consisted of periodic calls to debtors, with gradual escalation of the pressure applied.

Therefore, they were looking for a behavioral assessment tool that would fit into this process of ongoing phone calls with the debtor.

They realized that the debtors would not expose any high-risk behavioral tendencies voluntarily, so they wanted to evaluate behavioral speech analytics of their phone interactions as a possible source of insight.

They chose to evaluate Voicesense and their Debt Collection Predictor solution that provides a risk classification score (probability) of the behavioral tendency; in this case,

of potential high-risk behavioral tendencies based on analysis of the debtor's voice.

The evaluation objective was to assess the accuracy of the Voicesense debt collection predictive analysis.

The validation study consisted of 2669 phone interaction recordings with debtors of the collection agency. The recordings were taken from the initial conversation with these debtors.

All recordings were taken from a period of several months prior, so by the time of conducting the evaluation, the agency already knew whether each debt was collected successfully (debt paid) or unsuccessfully (defaulted debt). The agency labeled all the calls according to their default outcome (yes/no). The sample was designed such that about 50% of the collections ended with default, to enable proper training and testing of the predictive model.

Voicesense received the call recordings audio from the agency. According to predictive analytics best practices, the

sample was divided into separate training and testing sub samples.

The agency gave Voicesense the default labels (yes/no) only for the training sample.

Voicesense ran its vocal analysis with its machine learning modeling and trained the predictive model on the training sample calls. The model was then applied to the testing sample and Voicesense gave the agency the default prediction classification for the testing sample calls.

The company then compared Voicesense predictive classifications for the testing sample calls to the actual default outcome.

THE RESULTS

Voicesense classified the debtors into five risk prediction categories—Very Low, Low, Moderate, High, and Very High Risk. Each category consisted of about 20% of the debtors.

The results showed that the default rate among the debtors that were classified by Voicesense as “Very High Risk” was three times higher

than the default rate among the debtors that were classified by Voicesense as Very Low Risk. 38 out of 135 debtors defaulted on their loan collections in the Very Low Risk group, compared with 115 out of 137 debt collection defaults in the Very High Risk Group.