Credit Risk Analysis - Part 2 (Financial Benefits)

Submission

By Malastare AI

Team Members

Mathew Hart

Joey Spreitler

Rihad Variawa

About the Document

This Document focuses on the following aspects of the project and termed as ***Part2***

1.Financial Benefit of the project

Note : *This document is intended for business stakeholders.For technical aspect of the project which was used to arrive at the financial benefit please read the Part1 of the project document.*

Purpose

CredeX is a leading credit card provider that gets thousands of credit card applicants every year. But in the past few years, it has experienced an increase in credit loss. The CEO believes that the best strategy to mitigate credit risk is to ‘acquire the right customers’.

Objective:

To help CredX identify the right customers using predictive models.The aim of this document is to provide Financial benefits of using predictive models in making predictions related to which customer will default and which customer would not.

The financial benefits are carved based on the results from applying predictive models on the Credex dataset Vs conventional ‘without model’ technique

Summary of Technical Analysis

- The project was performed using the CRSIP DM framework.
- A standard exploratory data analysis was performed on the data provided to understand variable importance and data quality issue.
- Data cleaning and and its preparation for model building was done.
- A few different models such as Logistic regression and Random forest was built on the prepared dataset.
- A model which was built using Random First was chosen as the best and final model.
- The final model upon prediction gave an *accuracy* of **67.62%** with the *ability to identify good customers* at **67.79%** and ability to identify bad customers at **64.49%**
- A scorecard was built and a cutoff was determined to identify good vs bad customers.The cutoff identified was **370.**Anyone below the cut-off are highly likely to default and anyone above the cutoff are highly unlikely to default.

<u>Following are variables used in the final model and its order of importance,The top variable being the most important and the bottom one being the least important,</u>

1No_of_times_90_DPD_or_worse_in_last_6_months
2.No_of_times_30_DPD_or_worse_in_last_6_months
3.No_of_times_60_DPD_or_worse_in_last_6_months
4.No_of_times_90_DPD_or_worse_in_last_12_months
5.No_of_times_30_DPD_or_worse_in_last_12_months
6.No_of_times_60_DPD_or_worse_in_last_12_months
7.Avgas_CC_Utilization_in_last_12_months
8.No_of_PL_trades_opened_in_last_12_months
9.No_of_PL_trades_opened_in_last_6_months
10.No_of_trades_opened_in_last_12_months
11.No_of_Inquiries_in_last_12_months__excluding_home___auto_loans_
12.No_of_trades_opened_in_last_6_months
13.No_of_Inquiries_in_last_6_months__excluding_home___auto_loans_
14.Total_No_of_Trades
15.Outstanding_Balance

Financial Benefits

Confusion Matrix		Actual Output	
		Good Customer ("yes")	Bad Customer ("no")
Predicted Output	Good Customer ("yes")	45235	1251
	Bad Customer ("no")	21328	1687

The above table depicts data from a confusion matrix that was created on the Credex dataset (that excludes rejected population), which helps to determine the gain achieved using the ‘ With Model’ vs ‘No Model’ on the applicants who were provided with credit card.

- The model was run on a dataset that contains 69501 customers that does not include the rejected population
- The model was able to predict 45235 customers as good customers
- The model was able to predict 1687 customers as bad customers
- The model failed to predict 22579 customers into the correct classification

Given the above data the next slide explains the financial benefits obtained from the model Vs with our a model

Profit Calculation – ‘With model’ v/s ‘No Model’

Consideration - We are considering that on an average profit/loss occurred for each customer by credit card companies from variables such as ‘Interest’, ‘fees charged to cardholders’ and ‘Transaction fees paid by merchants’.Based on these assumptions below is the profit/loss used to calculate the financial benefit.

- Profit - ₹ 2500
- Loss - ₹ 50,000

Profit /Loss calculation on data classified without ‘Model’,i.e based on the original Credex dataset.The dataset has 66563 good customers and 2938 bad customers

- Profit = 2500* 66563
- Loss = 50000* 2938
- Net Profit = Profit - Loss

Profit/ Loss calculation on data classified using ‘Model’

- Profit = [2500* true positive(45235) + 50000* true negatives(1687)]
- Loss = [2500*False negative(21328)+50000*false positive(1251)]
- Net Profit = Profit - Loss

Net Financial gain

- Net Financial gain using the model = 81567500 - 19507500 = **₹ 6,20,60,000**
- Percentage financial gain = 6,20,60,000 / 1,95,07,500 *100 = **318.13%**

	Without Model	With Model
Profit	166407500	197437500
Loss	146900000	115870000
Net Profit	19507500	81567500

From above calculation it is clearly understood that there is a gain of 318% when a model is used to identify prospects and defaulters.Which is a clear and efficient way to easily identify prospects and defaulters.

Credit loss saved & Revenue loss

Credit Loss:

Credit loss occurred to the organization when the customer defaults, for the data that we have used, credit loss is about 4.22%. After building the model credit loss can be reduced to 2.42%.

Total Credit loss saved = 1.8%

Revenue Loss:

Revenue loss occurred to the organization when the good customers are wrongly identified as likely to default or bad customers by model.

Total Revenue loss = 32%

Conclusion :

Given the power of predictive modelling Credex can gain a profit by 300% compared to conventional method of identify defaulters/non-defaulters.Scores help the bank staffs to clearly understand the prospects and distribute loan amount accordingly and are able to easily decide whom to accept /reject as a customer.

END of Slide(Part 2)

Thank You