# Merchant Credit Worthiness

Manan Arora
Urvin Soneta
Neha Gupta
Shubham Saboo
Balbir Singh

28th November 2019

—

GCLS IV: Fin Tech

—

Gaurav Sharma

## Problem Statement

ABC Corp. is a merchant acquirer (the company that would signup merchants for accepting credit card. The merchants get paid by the merchant acquirers for the amount spend done on credit card).

It has been in the business for more than 3 years and has around 300K merchants on its platform. The revenue on acquiring side are going down due to the pressure of merchant discount rate (MDR).

The company is looking for lending to the merchants in its base as an opportunity to increase revenue and profitability.

The way repayments for the loans will happen through deduction from the payments which merchant acquirer makes to the merchant.

As a chief product officers for the company, we need to determine a strategy to identify credit worthy merchants.

# Data Points

## Firmographic Data

- Date of Onboarding
- Time since business Incorporation
- The industry of the merchant (Restaurant, Grocery Store, etc.)
- Address for the Merchant

## Transaction Data

- Merchant ID: Unique ID
- Date: Date of transaction
- Gross Amount: Total amount for which merchant was expected to be paid
- Charge Back: Amount which is charged back due to non-delivery or fraud
- Net Amount: Gross Amount – Charge back

## Clickstream / App Data

- Data captured from the online interaction of merchant with the ABC merchant portal and app

## Utilizing the Data:

### 1. Firmographic Data

a. Time since business incorporation: This can be used as an important parameter to determine feasibility of the business, longer the business has been operating, higher the chances of the business being sustainable and hence lower the risk. In the beginning of the lending operations, this can be used as a first level filter, for example only business that have been in operations for 2 years or longer can be considered for lending facility.

b. Date of Onboarding: The length of the transaction data that we have of the company will depend the day it was on boarded with us as a merchant. In the initial phase of the lending operations, we can focus on the merchants that we have had long standing relationship with rather than the ones that were on boarded very recently as we will have little transaction data to back their analysis. This can again serve as a filtering criteria – only vendors on boarded 6 months or prior to that to be considered.

c. Industry of the merchant: Industry in which the merchant operates can act as an extremely important factor. Some of things that can be considered here:

   i. Macro trends in the sector over the last couple of years (Growth, revenue, profits)
   ii. Volatility of the Sector over the last couple of years (share prices, competition)
   iii. Government regulation in the sector
   iv. Seasonality of the sector (Time series analysis might be needed)

d. Address data of the merchant: This data helps in inferring the size of market around the merchant and likely customer base – if it's a food joint, it's likely to be better off in a market place rather than near a residential society- clearly because of more footfall. Other data points can also be derived from this data, using open source tools like Google Places. Some of these data points are as mentioned below:

   i. Timings (Opening Hours): This can be used to understand the contact hours that the business has with its customers, it will vary widely according to the type of industry the merchant operates in. As an example, if we consider a pharmacy a 24 hours pharmacy will have a leverage over one which operates only 12 hours.
   ii. Vicinity: Through the address we can determine places existing in the vicinity of the merchant. This can help in two ways – one it will help determine density of a particular business types in an area, if the area is already crowded with businesses in the same space, it will make it ultracompetitive and hence margins will suffer on the other hand if there is lack of business in a particular area, it will make it easier for it to grow. Sometime existence of complimentary business in the vicinity also helps like a pharmacy near a hospital is more valuable as compared to pharmacy in a residential area.
   iii. Reviews and Ratings: Rating will provide an indication of customer satisfaction which will suggest the ability of the business to keep attracting customers and whether they will retain customers. This will help in verifying the business's ability to pay back a loan and its commitment and ability to carry out their business operations. The number of reviews along with their sentiment analysis would give us a holistic view of the legitimacy and quality of the company's business processes.

Since a business can exaggerate numbers and their business practices when asked, this information can be helpful to verify these details if a correlation exists. Additionally, this will enable us to track the change in customer sentiments.

## 2. Transaction Data:

a. Cash flow stability: The transaction data contains the net amount that we pay to the merchant each month. This can be used to see the stability in the cash flows a business generates over a longer period as stability is an important parameter. Trends on transaction data whether its static (constant each month), growing (at a constant or exponential rate) or dipping (either momentarily, cyclically or continuously) can be used to determine the consistency in the business operations which will be important in the loan repayment ability of any merchant.
b. Charge back ratio: Charge back is a negative factor as it's the amount of non –delivery at the merchant's end. This number is very important as it's in a way defaulter metric. The higher this amount, the worse it is. The ration of charge back with respect to the gross amount should not exceed 2% (industry standard)

## 3. Clickstream or App Data

a. Order Volumes: The app interaction data on the merchant app can be used to determine the order volumes of the merchant for specified period. This can be helpful in again factoring in the seasonality or consistency in the order books of the merchant.
b. Average Order Value: This can be used to determine the kind of customers the merchant is dealing with – generally bigger the average order value better it is. This can be used as a proxy to end customer trust in the merchant.

## Credit Worthiness Model

The way in which we can decide whether to give credit or not:

1. Initial filtering out of the existing merchants with us:
    a. Time of Business incorporation > 1.5 year (To ensure the business is market tested)
    b. Time of Onboarding > 1 year (To ensure we have sufficient cash flow records)
    c. Average monthly chargeback ratio is less than 2% (Industry standard)
2. Next we will cluster the remaining merchants based on the sector/ type of establishment (This data can come in both from the firmographic data as well as the Google Places API data) *Alternatively, we can also cluster using unsupervised learning where similar type of merchants will be automatically form different clusters.*
3. After forming clusters of merchants, we further analyze them based on monthly cash flows (available through the transaction data) – these need to exhibit stability over a longer term (say 2 to 3 years) and growth (over the shorter term say an year).
4. Once we are able to identify these good merchants, we analyze their secondary data which is available through click stream and Google Places.
    a. Through clickstream, we can get to know their order book diversification, ideally it shouldn't be a business that is only dependent on one or two products for their entire revenue and only limited repeat customers; there should be new customers as well.
    b. Through Google Places, we can analyze the rating and review of the place as well as the vicinity information.

Based on these merchants, we can define the sector wise metrics, their corresponding weights and final decision threshold for our model. This can be deployed for scrutinizing all the other merchants as well as the new incoming merchants for the lending business.

This rules based manual method can be utilized to prepare the labelled dataset for two years. Post this we can use this data along with defaulters list to train an automated machine learning model which will automatically do feature selection as well as weightages for these features.

Calculation of Lending Amount:

1. Once we have determined the feasibility of handing out a loan to vendor or not, we can then calculate their average monthly cash flow over the past year.
2. We can also keep the reimbursements given to them after deducting MDR.
3. Post that we can calculate the maximum loan amount that can be sanctioned as a multiple of the cash flow. It can be a multiple of 6 or 12 depending on the business size.