

NAME OF THE PROJECT

Micro-Credit Defaulter Model

Submitted by:

NEHA SHAHARWALE

**ACKNOWLEDGMENT**

Apart from the efforts of me, the success of any project depends largely on the encouragement and guidelines of many others. I take this opportunity to express my gratitude to the people who have been instrumental in the successful completion of this project.

I would like to show my greatest appreciation to Flip Robo and Data Trained. I can’t say thank you enough for his tremendous support and help. I feel motivated and encouraged every time I attend his meeting. Without his encouragement and guidance this project would not have materialized.

The guidance and support received from all the members who contributed and who are contributing to this project, was vital for the success of the project. I am grateful for their constant support and help.

**INTRODUCTION**

* Business Problem Framing

A Microfinance Institution (MFI) is an organization that offers financial services to low income populations. MFS becomes very useful when targeting especially the unbanked poor families living in remote areas with not much sources of income. The Microfinance services (MFS) provided by MFI are Group Loans, Agricultural Loans, Individual Business Loans and so on. Many microfinance institutions (MFI), experts and supporting the idea of using mobile financial services (MFS) which they feel are more convenient and efficient, and cost saving, than the traditional high-touch model used since long for the purpose of delivering microfinance services. Though, the MFI industry is primarily focusing on low income families and is very useful in such areas, the implementation of MFS has been uneven with both significant challenges and successes. Today, microfinance is widely accepted as a poverty-reduction tool, representing $70 billion in outstanding loans and a global outreach of 200 million clients. We are working with one such client that is in Telecom Industry. They are a fixed wireless telecommunications network provider. They have launched various products and have developed its business and organization based on the budget operator model, offering better products at Lower Prices to all value conscious customers through a strategy of disruptive innovation that focuses on the subscriber. They understand the importance of communication and how it affects a person’s life, thus, focusing on providing their services and products to low income families and poor customers that can help them in the need of hour.

Conceptual Background of the Domain Problem

Group Loans, Agricultural Loans, Individual Business Loans domains are the background essentials.

* Review of Literature

Microfinance is the provision on that provides access to  
various financial services such as credit, savings, micro  
insurance, remittances, leasing to low-income clients  
including consumers and the self employed, who  
traditionally lack access to banking and related services. Its  
main objective is to provision permanent access to  
appropriate financial services including insurance, savings,  
and fund transfer. It is rather an important tool for the  
eradication of poverty. Poor will be able to deal with  
emergency and also make significant investment  
expenditures. As microfinance becomes more widely  
accepted and moves into main stream, the supply of services  
to poor may also increase, improving the efficiency and  
outreach while lowering the costs. Microfinance must  
involve the people themselves in examining the problems and  
creating solutions if it is to be sustainable with those upon  
whom development is targeted.

* **Motivation for the Problem Undertaken**

Describe your objective behind to make this project, this domain and what is the motivation behind.

**Analytical Problem Framing**

* Mathematical/ Analytical Modeling of the Problem

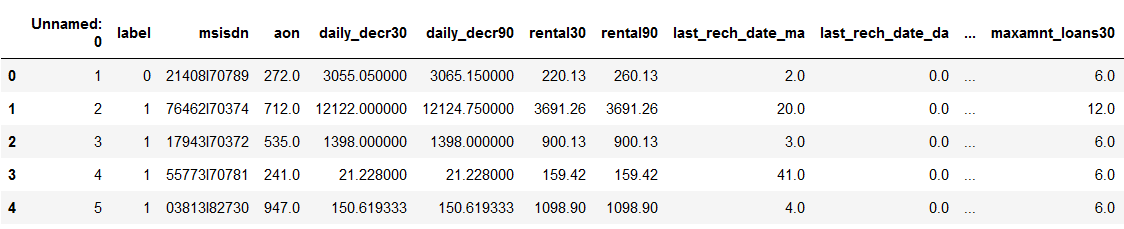
Describe the mathematical, statistical and analytics modelling done during this project along with the proper justification.

* Data Sources and their formats

We used a deep dataset of “MSBDataset” which includes the data

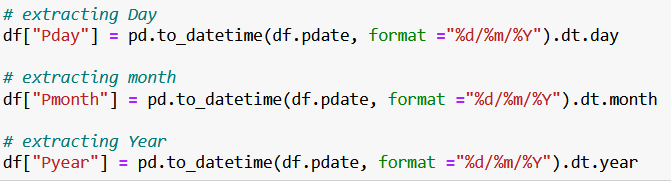
209593 columns and 37 features

All the data are in CSV and Excel. For reviews, the data looks like after adding in the data frame



* Data Preprocessing Done

1. At first, we need to train our model based on the data types .We checked the data types and all the data types were Integer and float, So there was no such requirement to change the data type .
2. There were no null values and missing values in the dataset
3. There was a data feature which we converted in day , month and Year column.



1. The three new columns were added in the data frame.
2. We checked for the outliers in the data set which was removed using Z score method.

Data Inputs- Logic- Output Relationships

There was hardly any correlation ship in the dataset so we used a method to select the best feature.

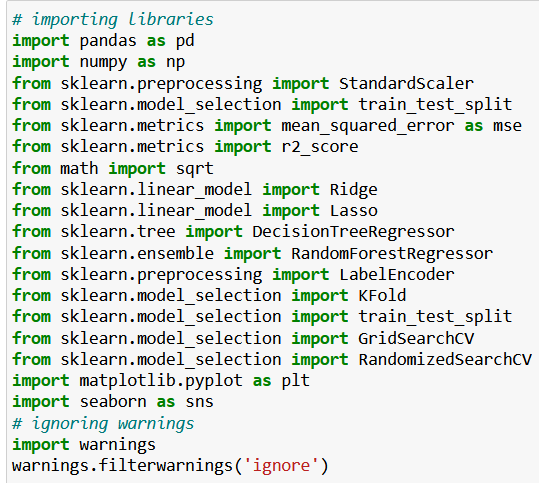
* State the set of assumptions (if any) related to the problem under consideration

After selecting the best feature we came up with the

good score.

* Hardware and Software Requirements and Tools Used

We used the Dataset, Laptop to complete the Project , The following were the libraries we added in the Project



**Model/s Development and Evaluation**

* Identification of possible problem-solving approaches (methods)

We used the Logistic regression model to achieve our project.

* Testing of Identified Approaches (Algorithms)

We used the Logistic regression model to achieve our project.

**CONCLUSION**

* Key Findings and Conclusions of the Study

We have implemented multiple layers of feature extractors and experimented with several  
classification algorithms to predict star rating from review text, which gives a good result.  
We carried out learning curve and ablative analysis, and experimented with difference feature  
extractors.

* Learning Outcomes of the Study in respect of Data Science

There is also further work that can be done in experimenting with the recommendation  
system. One option is to try more feature extraction approach such as bigram, trigram or  
word chunks. This could effectively increase the information that can be obtained from one  
review text, and add more features to overcome the high bias problem.