

Predicting Loan Defaulter: A Machine Learning Approach To Identify Bad Loans

Md. Abdullah Al Qayum

March 9, 2016

1 Introduction

Now a days the prediction of defaulting the borrower in future is a challenging task for credit card companies. Therefore the main objective of this project is to develop prediction models and a Data Product for defaulting the borrower in the future by taking advantage of available technological advancement.

The problem is to classify borrower as defaulter or non defaulter. It is commonly desired for banks to classify borrower accurately so as to manage their loan risk better and increase business .However developing such a model is a very challenging due to growing demand for loans.This project will attempt to use some traditional classification models like Logistic regression, Decision tree and Nave Bayesian classifier and attempt to create superior model using Sample Explore Modify Model and Assess (SEMMA) methodology to better predict customer default.And The final Outcome will be data product that will help end user to predict the customer behavior in Advance.

2 Targets

I will try to cover the concepts and tools that I will need throughout the entire data science pipeline , from asking the right kinds of questions to making inferences and publishing results. In this Project, I will try to apply all the skills I have learned by building a data product using real-world data.

- Understanding the Problem, and Getting the Data
- Exploratory Data Analysis and Modeling
- Prediction Model
- Creative Exploration:Apply Different Machine Learning Concept and Predict Accuracy
- Data Product