IMPACT OF ALTERNATIVE DATA AND TECHNICS IN BUILDING CREDIT SCORECARD: MERGED SCORING

KENNETH ASSOGBA

ABSTRACT. The aim of this work is to study impact of new data and technics in building credit scorecard. The final goal is to obtain a service that will evaluate creditworthiness of people not eligible for banking services and therefore improve their access to loans.

1. INTRODUCTION

Credit requests, of new and existing customers, are often evaluated by classical discrimination rules based on customers information. However, these kinds of strategies have serious limits. Many financial service providers see people on the sidelines as risky partly because they lack the data or other sources of information needed to identify high potential customers and their borrowing capacity.

2. PROBLEM

- Lack of traditional data to help build robust credit risk score system.
- Most credit risk systems are not designed keeping the rural segment in mind.
- Hard for financial institutions to build right combination of technology and talent.
- Hard to keep up with fast changing scientific and regulatory knowledge.

3. SOLUTION APPROCH

Utilises remote sensing data alongside a range of traditional and alternative data points to assess target population creditworthiness.

4. BENEFITS

4.1. To Lending Organization.

- Increase in approval rates
- Decrease in default rates
- Reduction in time it takes to process
- Increase automation and remove human errors

4.2. To Target Population.

- Increased access to loan
- Reduction in time it takes to process
- Build score which can help bring them into formal lending system for larger loans in future

5. Program

5.1. **Completed.** This section presents the work that was done in the first week.

- Review CGAP's nano credit literature
- Do some research on KYC
- Read articles on credit scoring

We have identified the data used in a classic scoring and the altinative data that we can introduce.

- Classic Data
  - Applicant Data
  - Credit Bureau Data
- Non-Traditional Data
  - Telecom Data
  - Social network Data
  - Mobile Data

On the similar projects that we have already been able to study, the target population was generally reduced and having common characteristics.
For example:

- Smallholder families
- Young graduate
- Young entrepreneur
- ...

So we wonder if we should not restrict the target population.

5.2. **Release.** This section presents the work that is going on.

- Analyse the existing lending data and study which alternative data sources, could improve the predictive power of lending decisions.
- Studies of the reports and data of the World Bank and the CGAP
  - CGAP Smallholder Families Data Hub
  - FY 2018 Benin Country Opinion Survey Report
  - Benin - Global Financial Inclusion (Global Findex) Database 2014
  - Benin - Global Financial Inclusion (Global Findex) Database 2018

5.3. **Backlog.** This section presents the work to be done.

- Analysis of existing solutions
- Comparison of techniques and algorithms used in scoring.
  - decision trees (Breiman et al., 1984)
  - neural networks (Mcculloch and Pitts, 1943)
  - discriminant analysis (Fisher, 1936; Mahalanobis, 1936)
  - logistic regression (Cox, 1970; Cox and Snell, 1989)
  - discriminant analysis and logistic regression
  - ...
- Tests
- Determine the tools to collect the data needed to train the model

References

[1] CGAP (Consultative Group to Assist the Poor): *CGAP and Harvesting Explore Use of Alternative Data in Credit Scores*,

cgap.org/news/cgap-and-harvesting-explore-use-alternative-data-credit-scores

[2] Farid Beninel, Waad Bouaguel, Ghazi Belmufti: *Transfer Learning Using Logistic Regression in Credit Scoring*

[3] Rory P. Bunker, M. Asif Naeem, Wenjun Zhang: *Improving a Credit Scoring Model by Incorporating Bank Statement Derived Features*

E-mail address: kennethassogba@gmail.com