

Competitive Machine learning : Fraudulent Online Loan Applications to Bitbond

O'JEANSON Baptiste

13 novembre 2015

1 Introduction

Bitbond is a startup offering to everyone the possibility to borrow money in affordable terms or to lend money with profitable interest rate. Bitbond operates thanks to a peer-to-peer Bitcoin lending platform. Thanks to their idea, small businesses get access to affordable loans while lenders earn profitable interest rates. Since Bitbond works exclusively with the digital currency Bitcoin, neither the borrower nor the lender need a bank account to participate in our global loan market.

Bitbond's mission is to make investing and financing globally accessible.

The main purpose of this project is to use the data of Bitbond (which are accessible to everyone) to analyse it and follow a machine learning procedure to identify the financial risks.

2 Problem setting

We can distinguish two approaches. The first one focuses on the borrowers and the second one on the loans. We have informations on the borrowers thanks to the registration process that occurs before the first loan application. We also have informations on the loan they ask for.

I am interested in predicting the risk, at the application time of the loan, that a person will default the terms of this precise loan (he is asking for) according to the his personal data (location, gender, employment, etc.), to his past experience of loan (if he has) and finally according to the similar past loans that other people asked for.

To estimate this risk, we use :

- The informations he gave on him at the application time, in case he never borrowed before. In that case we don't have any information on his past acts, so we can only focus on his details in comparison with others that were in the same conditions (the same profile asking the same kind of loan, what did they do?).
- The informations we have on his past acts, in case that is not his first time. First we look at his past history in terms of paying back and then we look at people who asked for the same kind of loan.

First, I would like to focus on the borrowers to determine some profiles. Then I would like to focus on the different loans funded and how did the things go after funding.

3 How Bitbond platform works

Anyone who wish to get a loan must satisfy some requirements :

1. He must register on the platform ;
2. He must inform about his name, his gender, his date of birth and his country of citizenship ;
3. He must inform about his phone, his address and if he had moved within the last 3 months ;
4. He must inform about his employment (self-employed, salaried, retired, etc.) ;
5. He must inform about different account he may have (ebay, paypal, facebook, etc.) to show his healthy activity ;

Then, Bitbond uses all those informations to compute indicators. In fact, those informations allow Bitbond to assess the ability of the borrower to pay back a loan : the creditworthiness (represents by a letter).

For this indicator, Bitbond asks for credit informations and compute a score. This score conducts the rating for the creditworthiness, as well as the type of employment, the connection of the personal PayPal account, the eBay feedback score, the country of residence and the payment history of bitcoin loans.

Write about how does the platform run.

4 Data

4.0.1 Borrower informations

4.0.2 Loan informations

5 First mining ideas

- Is there a link between the location and defaulting (and the funding) ?
- Does the difference between the amount requested and the amount funded is bigger according to the person's place, his employment, his incomes ?
- Does there is a link between the borrower rating and the amount funded ?
- Is there a difference of time (between the demand of loan and the actual funding) more important while the borrower rating is F than A ?

- Are all the defaulting loan actually fraudulent? (fraudulent label are shit)
- Is the nominal interest rate higher when the borrower rating is F?