

PROJECT MCNULTY PROPOSAL

CLASSIFYING KIVA LOANS DEFAULT RISK

-Hiranya Krishna Kumar

DOMAIN

In Project McNulty, I will use the dataset of microfinance site Kiva to classify Loan default risk. The data was sourced from Duke University contains 5000 data points and 42 features. It is a binary class classification problem to predict if the loan is delinquent or not. I can fetch additional data by making API calls to KIVA.

DATA

id: Loan ID

1. description.languages: Language of loan description with levels English (en), Spanish (es), French (fr), Indonesian (id), Russian (ru)
2. status: Loan status with levels defaulted, in_repayment, and paid, see <http://build.kiva.org/docs/data/loans> for what each level means
3. funded_amount: Amount of loan which has been purchased by Kiva lenders
4. paid_amount: Amount of the loan which has been paid off
5. activity: Activity for which loan is requested
6. sector: Sector for which loan is requested
7. location.country_code: Country code
8. location.country: Country name
9. location.town: Town name
10. location.geo.level: Indicator for whether the latitude and longitude given is for country or town
11. partner_id: Field partner ID for local lending institutions, see a complete list at <http://api.kivaws.org/v1/partners.json>
12. borrowers.first_name: First name of borrower
13. borrowers.last_name: Last name of borrower
14. borrowers.gender: Gender of borrower
15. borrowers.pictured: Whether or not the borrower provided a picture
16. terms.disbursal_amount: The amount of money distributed to the borrower in the local currency. Comparing this amount to the loan amount shows the currency conversion rate locked in for the loan when it was posted.
17. terms.disbursal_currency: The ISO 4217 code for the currency used to distribute the loan to the borrower. This is usually the local currency for the borrower's country.
18. terms.disbursal_date: The date at which the funds from the loan were given to the borrowers. Note that it is possible for the money to be disbursed to borrowers before the loan is posted on Kiva.

PROJECT MCNULTY PROPOSAL

CLASSIFYING KIVA LOANS DEFAULT RISK

-Hiranya Krishna Kumar

- 19. terms.loan_amount: The amount of money distributed to the borrower in lender's currency.
- 20. terms.loss_liability.nonpayment: Categorical variable indicating who is liable for loss on a loan due to nonpayment, lender or partner.
- 21. terms.loss_liability.currency_exchange: Categorical variable indicating who is liable for loss on a loan due to currency exchange, partner, shared, or none.
- 22. posted_date: Date when loan was posted on Kiva.
- 23. funded_date: Date when loan was fully funded on Kiva.
- 24. journal_totals.entries: Number of journal entries (updates on the Kiva website) by borrower.
- 25. paid_date: Date when loan was paid back in full by borrower.
- 26. defaulted_date: Date when borrower defaulted on a payment on the loan.
- 27. terms.local_payments.due_date: Date when payment is due to the field partner
- 28. terms.local_payments.amount: Amount (in local currency) due to the field partner
- 29. terms.scheduled_payments.due_date: Date when payment is due to the lender
- 30. terms.scheduled_payments.amount: Amount (in US dollars) due to lenders.
- 31. delinquent: Whether or not loan has become delinquent(TARGET)**
- 32. video.youtube_id: If a video was posted, its YouTube ID. To see video, go to "<http://www.youtube.com/watch?v=>" + "video.youtube_id".
- 33. basket_amount: Amount of the loan which lenders have saved in shopping baskets, but has not been confirmed as purchased.
- 34. amount: Payment amount in US dollars
- 35. payment_id: Payment ID
- 36. local_amount: Payment amount in local currency
- 37. processed_date: Date payment is processed
- 38. rounded_local_amount: Rounded local payment amount
- 39. settlement_date: Settlement date of payment
- 40. lat: Latitude of loan location, as indicated by location.geo.level
- 41. lon: Longitude of loan location, as indicated by location.geo.level