For this project we will be exploring publicly available data. The aim is to build a model that will predict whether or not the borrower paid back their loan in full.

We will use lending data from 2007-2010 from lending club.com.

Here are what the columns represent:

* credit.policy: 1 if the customer meets the credit underwriting criteria of LendingClub.com, and 0 otherwise.
* purpose: The purpose of the loan (takes values "credit\_card", "debt\_consolidation", "educational", "major\_purchase", "small\_business", and "all\_other").
* int.rate: The interest rate of the loan, as a proportion (a rate of 11% would be stored as 0.11). Borrowers judged by LendingClub.com to be more risky are assigned higher interest rates.
* installment: The monthly installments owed by the borrower if the loan is funded.
* log.annual.inc: The natural log of the self-reported annual income of the borrower.
* dti: The debt-to-income ratio of the borrower (amount of debt divided by annual income).
* fico: The FICO credit score of the borrower.
* days.with.cr.line: The number of days the borrower has had a credit line.
* revol.bal: The borrower's revolving balance (amount unpaid at the end of the credit card billing cycle).
* revol.util: The borrower's revolving line utilization rate (the amount of the credit line used relative to total credit available).
* inq.last.6mths: The borrower's number of inquiries by creditors in the last 6 months.
* delinq.2yrs: The number of times the borrower had been 30+ days past due on a payment in the past 2 years.
* pub.rec: The borrower's number of derogatory public records (bankruptcy filings, tax liens, or judgments).
* not.fully.paid: 0 means the borrow did not pay the back the loan in fill, 1 means he did. This is the target class.

**Tasks:**

Build a model using any algorithm of choice to predict whether or not the following 10 individuals will pay back their loans or not.

**PS: The model accuracy should not be less than 80%**

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **credit.policy** | **purpose** | **int.rate** | **installment** | **log.annual.inc** | **dti** | **fico** | **days.with.cr.line** | **revol.bal** | **revol.util** | **inq.last.6mths** | **delinq.2yrs** | **pub.rec** |
| 1 | debt\_consolidation | 0.1189 | 829.1 | 11.3504065 | 19.48 | 737 | 5639.95833 | 28854 | 52.1 | 0 | 0 | 0 |
| 1 | credit\_card | 0.1071 | 228.22 | 11.0821426 | 14.29 | 707 | 2760 | 33623 | 76.7 | 0 | 0 | 0 |
| 1 | debt\_consolidation | 0.1357 | 366.86 | 10.3734912 | 11.63 | 682 | 4710 | 3511 | 25.6 | 1 | 0 | 0 |
| 1 | debt\_consolidation | 0.1008 | 162.34 | 11.3504065 | 8.1 | 712 | 2699.95833 | 33667 | 73.2 | 1 | 0 | 0 |
| 1 | credit\_card | 0.1426 | 102.92 | 11.2997322 | 14.97 | 667 | 4066 | 4740 | 39.5 | 0 | 1 | 0 |
| 1 | credit\_card | 0.0788 | 125.13 | 11.9049676 | 16.98 | 727 | 6120.04167 | 50807 | 51 | 0 | 0 | 0 |
| 1 | debt\_consolidation | 0.1496 | 194.02 | 10.7144178 | 4 | 667 | 3180.04167 | 3839 | 76.8 | 0 | 0 | 1 |
| 1 | all\_other | 0.1114 | 131.22 | 11.0020998 | 11.08 | 722 | 5116 | 24220 | 68.6 | 0 | 0 | 0 |
| 1 | home\_improvement | 0.1134 | 87.19 | 11.407565 | 17.25 | 682 | 3989 | 69909 | 51.1 | 1 | 0 | 0 |
| 1 | debt\_consolidation | 0.1221 | 84.12 | 10.2035921 | 10 | 707 | 2730.04167 | 5630 | 23 | 1 | 0 | 0 |