

Test Questionnaire

Business Problem

Bank ABC wants to predict app score for current credit card customers. The app score will denote a customer's credit worthiness and help the bank in reducing credit default risk.

The target variable denoted as Bad_flag_worst6.

If value is 0 – it means customer has good credit history

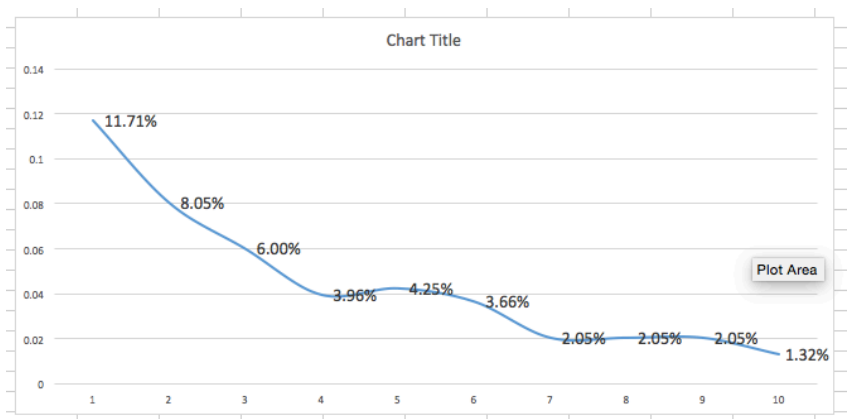
If value is 1 – it means customer has falls into 30 DPD + bucket

Tookitaki Benchmarks

We have build ensemble model with gini as 37.9. Below is our rank ordering

Rank Ordering

Decile Rank	66 features (Ensemble)
10	0.11713031
9	0.08052709
8	0.06002928
7	0.03958944
6	0.04245974
5	0.03660322
4	0.02052786
3	0.0204978
2	0.0204978
1	0.01317716



Data Explanation

- Account_Segment: customer's historical accounts data and payment behavior history
- Enquiry_Segment: customer's historical enquiry data such as enquiry amount, and enquiry purpose
- Data: current customer applications with demographic data
- Data given for April-Dec 2015.

Only 4.29% of the entire data falls under Bad_flag_worst6.

We are sharing 2 splits of data – 70% and 30%

Test Questionnaire

ASK: You need to build a model on 70% data and test on 30% data and share the following with us over word doc –

- Data exploration insights – what did you find and what decision did you take?
- Feature matrix - List of features selected with gain (kindly share gain calculation) and how you calculated them
- Model evaluation - Gini and rank ordering

Please submit the code in github.

Some feature suggestions

Feature	Gain	Remarks
payment_history_avg_dpd_0_29_bucket	0.045653821	mean count of accounts that is in 0-29 dpd bucket throughout the payment history
total_diff_lastpaymt_opened_dt	0.04083407	The total duration between last payment date and account opened date of all accounts
min_months_last_30_plus	0.038216344	The smallest number of months passed before first 30+ dpd appeared for each account.
utilisation_trend	0.037523553	[total cur_bal_amt / total credit limit] / [mean cur_bal_amt / (mean credit limit+ mean_cashlimit)]
count_enquiry_recency_365	0.036216624	# of enquiry made in past 365 days
Ratio_currbalance_creditlimit	0.035640536	[total cur_bal_amt / total credit limit]
mean_diff_lastpaymt_opened_dt	0.032185579	The average duration between last payment date and account opened date of all accounts
mean_diff_open_enquiry_dt	0.031057926	average difference between enquiry dt_opened date and enquiry date
payment_history_mean_length	0.029258469	average length of payment_history variable
max_freq_enquiry	0.024501434	most frequent enquiry purpose
count_enquiry_recency_90	0.023796745	# of enquiry made in past 90 days
perc_unsecured_others	0.023428297	ratio of secured loan type enquiry purpose to total enquiry purpose made