Application Credit Risk Score Card – Home Loans

Introduction:

The report explains the end-to-end process of building a score card for approving or rejecting a home loan application. To ensure model interpretability, logistic regression was used to predict the probability of default and Weight of Evidence (WOE) encoding was applied to the data.

The project follows standard industry practice of collecting data, identifying the target variable, creating and selecting independent variable, and finally developing a model.

Author Note:

This project was independently designed and implemented by Abhijit Kudtarkar as part of advanced credit risk modeling practice. Generative AI tools were used selectively for ideation support and documentation refinement.

Step 1: Data Generation:

As the data used in the banks is not present in the open source in same format and also in terms of volume. The following datasets were generated artificially:

1. **Approved Application Data:**
   A total of **100 thousand** home loan application records were generated through python code. The data included features like credit score, loan to value ratio, income to debt ratio, and other relevant variables.

2. **Bureau Tradeline Data:**
   For the approved home loan applicants, bureau data **at tradeline level** were synthesized. This included features like loan type, loan amount, balance, tenure, age of the loan and delinquency indicators in the last 3,6 and12 months the time of application.

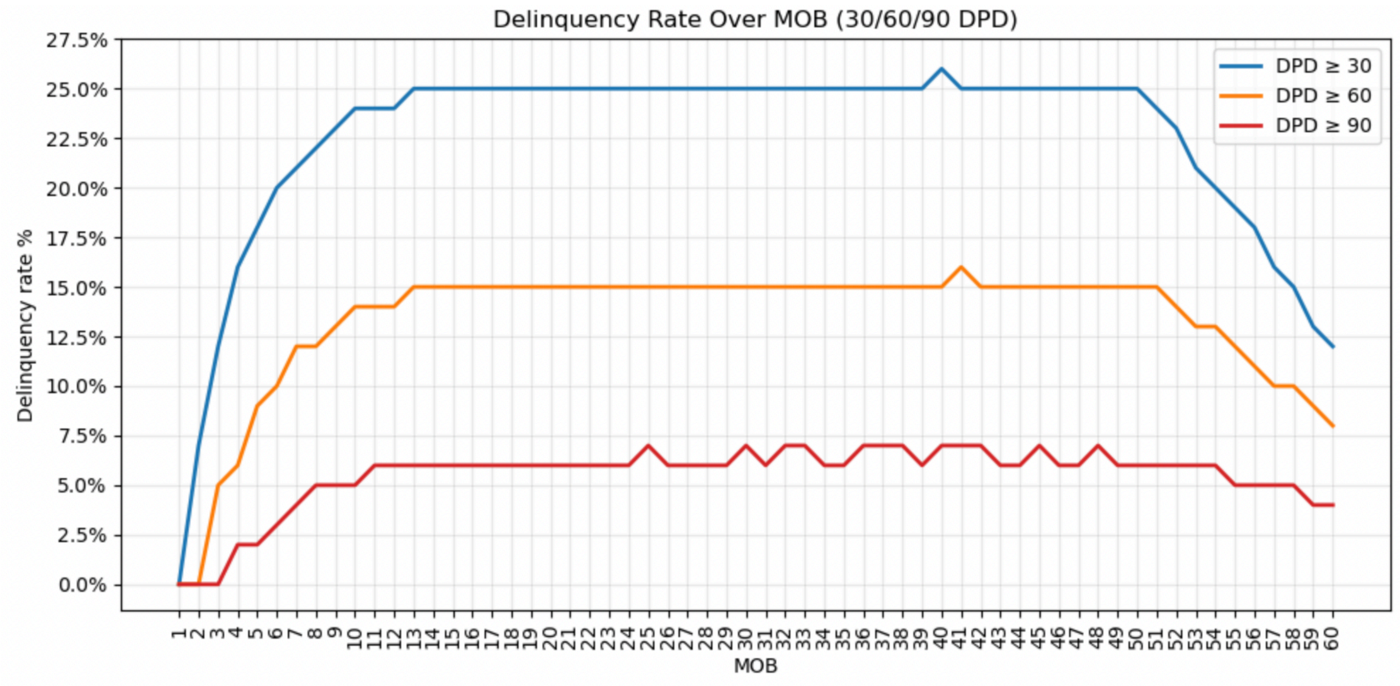
3. **Bureau Inquiry Data:**
   Bureau inquiry data is similar to bureau tradeline data. This dataset is at **inquiry level** and includes features like loan type, requested loan amount and indicators of whether the inquiry was made in last 3, 6 and 12 months at the time of application

4. **Post-approval Loan Performance Data:**
   Post the loan was approved, the month of month performance of these loans were generated which contained features like current DPD, outstanding balance, amount paid and months on book (MOB).

Step 2: Defining and Creating the Target Variable:

In this step, indicators were created to identify whether a loan defaulted at 30+,60+ and 90+ days past due across different months on book (MOB). The data was then aggregated by MOB to calculate the total number of loans, as well as the percentage of total loans in 30+, 60+ and 90+ delinquent. A vintage chart was created to decide the target variable

In the below vintage chart, there is a sharp increase in %30+, %60+ and %90+ delinquency till 11 MOB. The %90+ delinquency remains steady from 12 MOB onwards, and from 50 MOB it declines gradually. Therefore, **90+ delinquency at 12 MOB** was selected as **the target variable** for model development.



Step 3: Feature Creation:

A total of 931 features were created from bureau tradeline and bureau inquiry data to capture customer credit card exposure, repayment behavior, and credit-seeking patterns. After merging these with the application data, a total of **947 variables** were available for model selection.

Step 4: Feature Selection:

As the data was synthesized, the data was split into train and test using stratified random sampling with a ratio of 70:30. Therefore, 70 thousand records were used for feature selection and model building and the remaining 30 thousand were used for testing the final model

This step is performed to reduce total number of independent variables used for scorecard modelling

I. **Removing columns with single value or near-constant behavior (99.5% same value):**
   None of the **947** variables satisfied this condition so no columns were dropped here.

II. **Weight of Evidence encoding and IV screening:**
   Weight of Evidence (WOE) encoding was applied to all 947 variables. WOE helps handle outliers and create a linear relationship between the predictors and default. Prior to WOE transformation, coarse binning was performed to ensure that numerical variable showed a monotonic relationship with WOE.
   After WOE encoding, Information Value (IV) was calculated for all WOE encoded variables. Variables with IV>=0.02 and IV<=0.6 were considered for the next step. As a result of this step, **248** WOE encoded variables were selected out of 947.

III. **Correlation of WOE encoded variables:**
   Correlation measures how strongly two variables move together. If a pair of variables have absolute correlation value greater 0.85. Then one with least IV is dropped.
   After this step, the total variables for model building came down to **81** from 248.

IV. Variance Inflation Factor (VIF):

VIF (Variance Inflation Factor) tells how much a variable is being explained by other variables in a model
In this step of VIF of all variables is calculated and if a variable has VIF >10, then the variable is dropped and VIF is calculated for all other variables until each of them have VIF <=10
After this step, **70** variables were selected from the remaining 81 variables.

V. L1 Regularization:

L1 Regularization adds penalty and shrinks coefficients of the weak predictors to zero. Thus, acting as an automatic feature selection.
After this steps, **32** variables were selected for model building.

Step 5: Model Development and Performance:

I. Model Development:

Logistic regression is chosen as the final model for the credit scorecard because it provides probability of default as its output. When input variables are transformed using Weight of Evidence (WOE), the relationship between risk and predictors becomes linear. The coefficients of this model are easy interpret and thus will not be considered a black box model like deep learning models.
The appendix contains the final list of variables used in the model, along with their descriptions and model coefficients

II. Model Performance:

Metrics	Value	Interpretation
AUC	0.68	An AUC of 0.68 means the model does a decent job of separating risky borrowers from safer ones but not perfect. It is clearly better than random and usable for decision making
Gini	0.36	Gini = 0.36 means the model is decent at separating risky borrowers from safe one. It is better than random but not very strong
KS Statistic	0.28	KS = 0.28 means the model has moderate separation between risky and sage customers (about 28% max gap in their score distributions. It is acceptable but not strong for a scorecard.

III. Decile Ranking

	Training Data Set			Testing Data Set		
decile	**Total loans**	**Total loans 90+ at 18 MOB**	**%90+ at 18 MOB**	**Total loans**	**Total loans 90+ at 18 MOB**	**%90+ at 18MOB**
1	7,009	276	3.90%	2,926	110	3.80%
2	6,991	319	4.60%	2,963	144	4.90%
3	7,000	361	5.20%	2,994	155	5.20%
4	7,073	402	5.70%	3,091	183	5.90%
5	7,046	452	6.40%	3,124	186	6.00%
6	6,881	483	7.00%	2,996	185	6.20%
7	7,000	555	7.90%	2,947	229	7.80%
8	7,000	615	8.80%	3,003	242	8.10%
9	7,000	869	12.40%	3,019	360	11.90%
10	7,000	1,699	24.30%	2,937	790	26.90%
Total	**70,000**	**6,031**	**8.62%**	**30,000**	**2,584**	**8.61**

	Training Data Set			Testing Data Set		
	Total loans	**Total loans 90+ at 18 MOB**	**%90+ at 18 MOB**	**Total loans**	**Total loans 90+ at 18 MOB**	**%90+ at 18 MOB**
Total	70,000	6,031	8.62%	30,000	2,584	8.61%
If loans in Decile 1 to 7 are approved	**49,000**	**2,848**	**5.81%**	**21,041**	**1,192**	**5.67%**
% Drop in the book	30%	53%	2.80%	30%	54%	2.95%

Conclusion:

The decile cut-offs were derived from the predicted probability of default in training data and then applied to the testing data, ensuring no data leakage. A clear and monotonic increase in 90+ delinquency at 18 MOB across deciles in both training and testing data demonstrates good risking ranking and stable behaviors.

Approving loans in deciles from 1to 7 removes about 30% of the loans in the testing data while reducing the 90+ delinquency at 18 MOB rates from 8.61% to 5.67% which means 54% of bad loans will be removed if we remove 30% of the portfolio at the time of application.

As the data used for training and testing the model was synthesized, the observed AUC of 0.68, Gini of 0.36, and KS of 0.28 reflect moderate separation. A stronger performance can be expected when the model is trained and tested on the real-world data.

Appendix

Sr No.	Variables	Description	Model Coefficient
1	OFF_US_Property Loan_DPD30p_L3M_sum_woe	Number of 30+ DPD events on off-us property loans in the last 3 months at the time of application	0.046852
2	OFF_US_Personal Loan_Inquiry_Requested_Amount_woe	Amount requested through personal loan inquiries with other lenders at the time of application	0.012495
3	OFF_US_Auto Loan_DPD30p_L3M_sum_woe	30+ DPD events on off-us auto loans in the last 3 months at application	0.043559
4	Loan_Amount_woe	Loan amount applied for at the time of application	0.568403
5	Property Loan_DPD30p_L12M_sum_woe	30+ DPD history on property loans in the last 12 months at application	0.085311
6	Auto Loan_DPD60p_L6M_sum_woe	60+ DPD events on auto loans in the last 6 months at application	0.086461
7	OFF_US_Credit Card_Inquiry_Requested_Amount_mean_woe	Average amount requested through credit card inquiries with other lenders at application	0.122078
8	Personal Loan_DPD60p_L6M_sum_woe	60+ DPD events on personal loans in the last 6 months at application	0.001258
9	Personal Loan_Enq_L30D_sum_woe	Number of personal loan inquiries in the last 30 days at application	0.086965
10	All_DPD30p_L3M_max_woe	Worst 30+ DPD across all loans in the last 3 months at application	0.085235

11	OFF_US_Credit Card_Enq_L30D_max_woe	Maximum recent credit card inquiries with other lenders in last 30 days	0.05043
12	All_DPD90p_L12M_sum_woe	Total 90+ DPD events across all loans in the last 12 months at application	0.014404
13	Auto Loan_DPD60p_L3M_sum_woe	60+ DPD events on auto loans in the last 3 months at application	0.010593
14	Credit Card_DPD60p_L3M_sum_woe	60+ DPD events on credit cards in the last 3 months at application	0.080424
15	Home Loan_DPD60p_L3M_sum_woe	60+ DPD events on home loans in the last 3 months at application	0.039067
16	OFF_US_DPD60p_L12M_sum_woe	60+ DPD events with other lenders in the last 12 months at application	0.029859
17	Personal Loan_DPD60p_L3M_sum_woe	60+ DPD events on personal loans in the last 3 months at application	0.035221
18	OFF_US_DPD90p_L6M_sum_woe	Severe delinquency (90+) with other lenders in the last 6 months at application	0.150975
19	All_Original_Loan_Amount_max_woe	Maximum original loan amount across all loans at application	0.085885
20	OFF_US_Credit Card_Balance_max_woe	Highest credit card balance with other lenders at application	0.053886
21	Personal Loan_DPD30p_L12M_max_woe	Worst 30+ DPD on personal loans in the last 12 months at application	0.088167
22	Home Loan_DPD60p_L6M_max_woe	Maximum 60+ DPD on home loans in the last 6 months at application	0.024235
23	OFF_US_Home Loan_DPD30p_L3M_sum_woe	30+ DPD events on off-us home loans in the last 3 months at application	0.032924
24	OFF_US_Auto Loan_Enq_L30D_sum_woe	Auto loan inquiries with other lenders in the last 30 days at application	0.043653
25	OFF_US_Personal Loan_DPD30p_L3M_max_woe	Worst 30+ DPD on off-us personal loans in the last 3 months at application	0.020979
26	Credit Card_Balance_mean_woe	Average credit card balance at the time of application	0.032181
27	ON_US_Credit Card_Enq_L30D_sum_woe	Credit card inquiries with the bank in the last 30 days at application	0.1254
28	Credit_Score_woe	Overall borrower credit quality at the time of application	0.812042
29	All_Enq_L30D_sum_woe	Total credit inquiries across all products in the last 30 days at application	0.094085
30	Cash_Reserves_Amount_woe	Cash reserves available at the time of application	0.589813
31	ON_US_Loan_Count_woe	Number of existing loans with the bank at application	0.064324
32	Debt_to_Income_Ratio_woe	Borrower debt-to-income ratio at the time of application	0.295479