Summary Report on Credit Scorecard Creation

Course Name:
IS453 - Financial Analytics

Class & Team No.:
G1T9

Team Members:
Edwin Tok, Elvin Ng, Jiaen Goh, Jolene Loh,
Phoenikelly Yong, Lim Wei Zhi

1 Business Opportunity & Our Hypothesis

Our team was hired as data analytics consultants to assist a start-up Fintech lender in developing a credit scorecard for use in evaluating credit applicants. Credit scorecards enable Fintech companies to forecast their customers' ability to repay their debts on time. We have identified potential credit applicants that the Fintech company would be well suited to lend to, and we have developed a credit scorecard that focuses on the variable of home ownership, which we believe is a significant factor in determining potential defaulters.

We hypothesised that homeowners would have a higher chance of not defaulting as compared to non-homeowners because homeowners can use their housing assets as collateral in the event of a default. To test our hypothesis, we will create a scorecard to see how homeownership affects credit scoring. We can then evaluate our hypothesis by analysing the score allocation between different attributes in home ownership.

2 Dataset Explanation & Analytic Techniques Used with Limitations

The Credit Risk Dataset was used for this analysis because it contained simulated credit bureau data that was available on Kaggle. There were 32581 rows and 12 columns consisting of both numerical and categorical variables.

2.1 Features of Dataset

- person_age (Age)
- person_income (Annual Income)
- peronhomeownership (Home ownership)
- Personemplength (Employment Length in years)
- Loan_intent (loan intent)
- Loan_grade (loan grade)
- Loan_amnt (Loan amount)
- Loanintrate (Loan interest rate)
- Loan_status (Loan status)
- Loanpercentincome (Percent income)
- Cbpersondefaultonfile (Historcial default)
- Cbpersoncredhistlength (Credit history length)

2.2 Analytics Techniques

We used Credit Scorecard, which helps to reduce the rate of default by determining a probabilistic threshold that can be used to decide the risk tolerance. It uses Weight of Evidence to establish the monotonicity, supports outliers and help to deal with missing value. It also uses information value which is used for univariate screening to reduce the number of independent variables.

We also used logistic regression to estimate the probability of whether an event will occur.

3 Data Preparation & Feature Selection

These are the steps that we took to prepare our data:

1. Dropping outliers
   a. Age >= 123 & Employment age >= 123
2. Filling up NA values with its mean
3. Remove highly correlated independent variables

a. cb_person_cred_hist_length and person_age has a correlation of 0.88. Hence, we dropped person_age.

4 Grouping, Screening & Scorecard Construction Approach

We opted to use Scorecardpy due to its simplicity in performing all steps of scorecard construction, such as optimal WOE binning and evaluating scorecard performance via AUROC and KS score.

4.1 WOE Binning

By using Scorecardpy's woebin and woebin_plot methods, we can generate and plot the WOE binnings for all independent variables, allowing us to identify non-monotonic and low IV independent variables.

We removed independent variables with IV < 0.02 due to its negligible predictive power and "cb_person_cred_hist_length" with IV = 0.0043. Then we discovered that the "loan_amnt" and "loan_int_rate" were non-monotonic, thus we manually adjusted the binning breaks until the trend was monotonic.

We also realised that some independent variables had IV > 0.5, indicating overfitting; however, due to our already low number of variables to consider, we decided to keep these variables to ensure our scorecard is comprehensive in predicting most characteristics.

4.2 Model Training

We proceeded to split the dataset into a train and test set for our logistic regression model using Scorecardpy's split_df method, based on a 3:7 ratio respectively. Using Scorecardpy's woebin_ply method, we did WOE encoding for both the train and test set, with the WOE bins generated from the previous step. We then initialized the X and y train and test variables, using the X and y train variables to build our logistic regression model using sklearn's LogisticRegression class and its fit method.

4.3 Model Evaluation

Using sklearn's LogisticRegression class and its predict_proba method with the X train and test variables, we calculated the probability estimates for the train and test set. We used Scorecardpy's perf_eva method to plot the KS and ROC curve to evaluate our models' performance with the y train and test variables, and the previous probability estimates.

We discovered that the KS score was 0.6379 and AUROC was 0.8861 for the train set, and the KS score was 0.6023 and AUROC was 0.8702 for the test set (refer to 9.3 in appendix). Given that the KS score fell between 0.4 - 0.7 and AUROC was between 0.8 - 0.9, we determined that our model performed well. We also used sklearn's confusion_matrix and classification_report method to create a confusion matrix and classification report to generate the classification metrics for further evaluation (refer to 9.4 in the appendix).

4.4 Scorecard Creation & Scaling (Refer to 9.1)

Lastly, we created the scorecard using Scorecardpy's scorecard method, with the WOE bins and the logistic regression model generated earlier. Using the pprint method, we printed the scorecard in an appropriate format. Then, using the bin, logistic regression, and columns from our independent variables, we scaled the scorecard with base points, where basepoints_eq0 is 0, points0 is 600, odds0 is 1/50 and pdo equals 20. These are the target odds for our scorecard for a rough 25% chance of an applicant being rejected.

5 Scorecard Walkthrough

5.1 Distribution of Points From Scorecard (Refer to 9.2)

Earlier, we generated WOE and IV for all our characteristics respectively and eliminated those with little predictive power to determine the probability of default. The table in the appendix shows our scorecard along with our final selection of characteristics. Before going through scoring examples, we take a closer look at how the points are allocated with each characteristic:

- Person_home_ownership: For home ownership, applicants that own a property or currently have a mortgage to one will have more points allocated. Those that rent apartments or otherwise receive demerit points.

- Person_income: For annual income, applicants need to have an amount greater than $60000 to be allocated more points. Those that fall under $35000 to $60000 receive no points, while those that fall below this bracket receive demerit points. The greater the income, the better the score.

- Loan_intent: For loan intent, business ventures, education and personal loans have a positive score allocated. On the other hand, home improvement, medical and debt consolidation loans are allocated demerit points.

- Loan_percent_income: For loan per cent income, values greater than 0.31 are allocated a large number of demerit points, whereas those below this value are allocated a positive score. The lower the number, the better the score.

- Person_emp_length: For the number of years a person is employed, applicants that work for less than three years are allocated three demerit points. For three to five years, no points are allocated. Applicants with more than five years receive three points. The greater the number of years, the better the score.

- Cb_person_default_on_file: This characteristic has the least weightage in this table due to the number of points allocated in either the ‘Y’ or ‘N’ case. In comparison to other characteristics, it might be its lesser predictive power that explains this observation.

- Loan_amnt: For loan amounts, amounts greater than $13000 are allocated one to two demerit points, amounts between $8000 to $13000 are allocated no points, while amounts less than $8000 receive one point. The lesser the amount, the better the score.

- Loan_grade: For loan grades, grades A to C have positive points allocated to them. Grades D and below are allocated 57 demerit points. The better the grade, the better the score.

- Loan_int_rate: For loan interest rates, percentage points greater than 14 per cent are allocated demerit points while those between 12 to 14 per cent are not allocated points. Finally, percentage points less than 12 per cent are given three points. The lower the interest rate, the better the score.

5.2 Determining Cut-off Score

We chose the cut-off score based on our assessment of the likelihood of a loan default being accepted for approval. We settled on a score that predicts roughly a 75% chance of acceptance. As we also scaled our scorecard with basepoints to help the business determine adverse action responses, a score of 527.0 will predict a rough 25% probability of being rejected.

5.3 Observations from Scorecard Point Allocation

While it is clear that positive points and demerit points are allocated based on certain attributes, some characteristics are allocated more points than others. Some of the observations we have made can help deduce the relative importance of predicting an applicant's probability of default. For example, person_home_ownership, person_income, loan_percent_income, loan_intent and loan_grade stand out amongst the other characteristics as they have a wide range of score allocations.

5.4 Scorecard Scoring Examples (Refer to 9.2)

In the appendix, we used examples of two applicants — one is a homeowner and the other is a renter — to see how the scorecard tabulates the total scores. In order to test our hypothesis, we kept the score allocations for all characteristics except for person_home_ownership to be the same. In total, the homeowner receives a score of 542.0 while the renter receives a score of 493.0. Bearing in mind that the cut-off score is 527.0, this means that if both applicants were to apply for a loan with largely similar applicant information, only one of them would attain a score higher than the cut-off. In this case, the homeowner passes the cut-off score and will be approved for credit. On the other hand, the renter fails to meet the cut-off score and will be denied credit.

6 Results of Scorecard and Tuning & Our Hypothesis Evaluation

From the scorecard scoring examples above, when all other characteristics score the same points, the homeowner scores higher than the renter. Regarding the approval process for a loan application, the homeowner will be approved while the renter will be denied.

We began our investigation with a hypothesis to test — homeowners would have a higher chance of not defaulting as compared to non-homeowners. When we look at the point allocation between the attributes for the person_home_ownership characteristic, owning a home sets an applicant 49 points higher than those that rent or do not have a home otherwise. Among the other characteristics across the scorecard, it also has the third-largest score point range. Given that home ownership dominates this score, this hypothesis has been proven true by our final scorecard.

7 Potential Issues with Scorecard with Solutions & Limitations

7.1 Scorecard

We believed that based on the population stability index, there could be changes in the population. Thus, the model would become less effective over time. We recommend that the credit scorecard be updated at least quarterly to mitigate this, bearing in mind changes in economic conditions that could affect housing and rent prices, population, or product behaviour In our scorecard, the credit score only uses financial factors for approving a loan. We believe that a more holistic approach should be taken when approving loans. Other factors include the education level.

7.2 Regression Model

It is difficult to obtain complex relationships using the logistic regression model. As the start-ups progress and grow, we recommend that you consider using more advanced machine learning models, such as Neural Networks.

7.3 Conclusion

To conclude, we suggest that the company develop a new scorecard to assess renters more fairly for strategic reasons. As proven by our hypothesis, renters will score lower than homeowners on our scorecard. The score allocation disparity indicates an assumption that renters have a higher chance of defaulting as the property is a liability rather than an asset that can be used as collateral.

However, this can be an overgeneralization of the renter population. Non-homeowners may have other assets, such as liquid cash, and choose not to buy a home for personal reasons, such as living with their parents. Therefore, it is unwise to reject a certain customer segment just because they do not own a home. The company could consider alternatives if they feel that is risky. They could lower the credit limit or raise the interest rate, thereby balancing out the risk and reward. By creating a new scorecard, the company can target a new customer segment and capture a new group of customers. Overall, they will be able to increase their reach and profits.

8 Challenges Encountered & Lessons Learnt from Our Analysis

This project was very insightful and practical as we could apply the concepts we learned. The idea of creating a credit scorecard felt relevant and exciting, as we were very interested in the approval process for loans. Solving problems in the financial industry and developing a scorecard, helped us build a deep understanding of the industry standards and techniques that are used. The project also helped us develop the ability to analyse financial data and reinforce what we learned in the labs.

The process of building the scorecard was difficult as we struggled with creating the business context and finding a relevant dataset to use for the project. Most of the datasets we discovered contained several rows of noisy data and did not correspond to the business context. Other datasets were lacking because we felt they lacked the information we required. We also had difficulty reaching an agreement on the business context and the creation of the scorecard. We debated which aspects of the business could provide more value to the company and generate more profits for the business. We eventually overcome the difficulties by weighing the benefits and value created for the company. We began working on the scorecard after we had established the business context. The process of creating the scorecard was relatively simple, and we did not encounter many difficulties because we had been sufficiently prepared by the labs from class.

To conclude, we learned the importance of developing a strong case for the business context and how the business context can help the company generate more revenue from a new customer segment. We were able to create a product that helps improve the company's business process by using advanced analytics. The scorecard will enable the company to quickly approve and reject more loans in a short period of time.

9 Appendix

9.1 Our Credit Scorecard

Characteristic	Attribute	Scaled w/BP
person_home_ownership	OWN	35.0
	MORTGAGE	19.0
	OTHER, RENT	-14.0
person_income	<35000	-24.0
	35000-60000	0.0
	60000-80000	10.0
	>80000	23.0
loan_intent	VENTURE	20.0
	EDUCATION, PERSONAL	9.0
	HOME IMPROVEMENT, MEDICAL, DEBT CONSOLIDATION	-12.0
loan_percent_income	<0.16	22.0
	0.16-0.31	3.0
	>0.31	-66.0
person_emp_length	<3.0	-3.0
	3.0-5.0	0.0
	>5.0	3.0
cb_person_default_on_file	N	0.0
	Y	1.0
loan_amnt	<8000	1.0
	8000-13000	0.0
	13000-18500	-1.0
	>18500	-2.0
loan_grade	A	31.0
	B	12.0
	C	2.0
	D,E,F,G	-57.0

loan_int_rate	<12.0	3.0
	12.0-14.0	0.0
	14.0-15.5	-6.0
	>15.5	-9.0

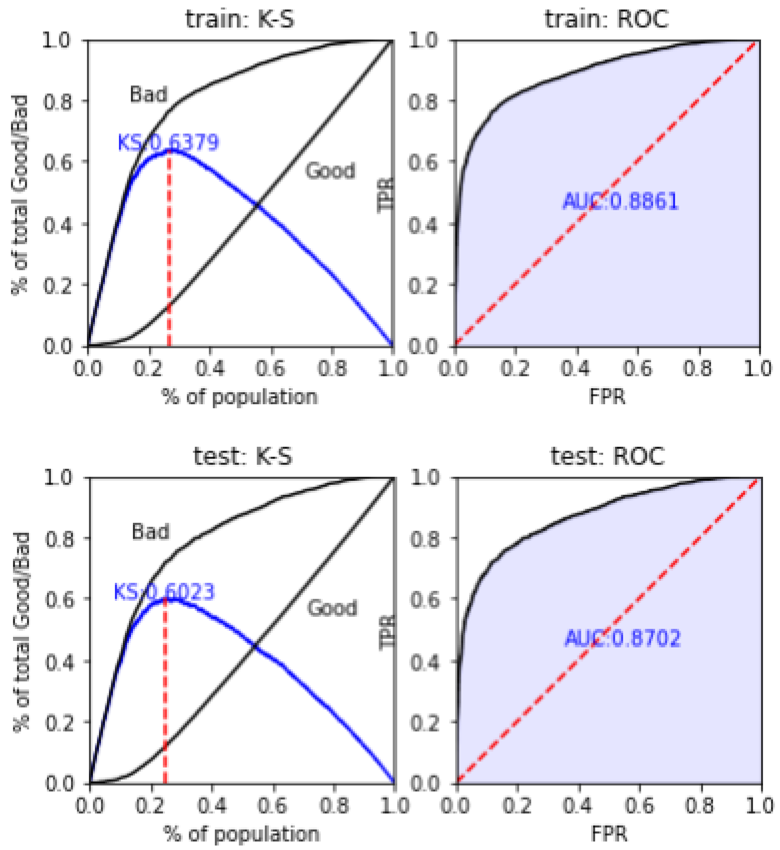
9.2 Scorecard calculation examples

Example 1 (Homeowner)	
Base points	527.0
Cb_person_default_on_file = N	0.0
loan_amnt = 30000	-2.0
loan_grade = D	-57.0
loan_int_rate = 10.65	3.0
loan_intent = DEBTCONSOLIDATION	-12.0
loan_percent_income = 0.06	22.0
person_emp_length = 7.0	3.0
person_home_ownership = OWN	35.0
person_income = 500000	23.0
TOTAL:	**542.0**

Example 2 (Renter)	
Base points	527.0
Cb_person_default_on_file = N	0.0
loan_amnt = 30000	-2.0
loan_grade = D	-57.0
loan_int_rate = 11.25	3
loan_intent = DEBTCONSOLIDATION	-12.0
loan_percent_income = 0.12	22.0
person_emp_length = 6.0	3.0
person_home_ownership = RENTER	-14
person_income = 550000	23.0
TOTAL:	**493.0**

Characteristic	Attribute	Scaled w/BP
person_home_ownership	OWN	35.0
	MORTGAGE	19.0
	OTHER, RENT	-14.0

9.3 KS and ROC Curve



9.4 Classification Report and Confusion Matrix

```
              precision    recall  f1-score   support

           0       0.95      0.89      0.92      8146
           1       0.59      0.77      0.66      1626

    accuracy                           0.87      9772
   macro avg       0.77      0.83      0.79      9772
weighted avg       0.89      0.87      0.88      9772

[[7262  884]
 [ 378 1248]]
```

9.5 Features and Description

Feature Name	Description
person_age	Age
person_income	Annual Income

person*home*ownership	Home ownership
person*emp*length	Employment length (in years)
loan_intent	Loan intent
loan_grade	Loan grade
loan_amnt	Loan amount
loan*int*rate	Interest rate
loan_status	Loan status
loan*percent*income	Percent income
cb*person*default*on*file	Historical default
cb*person*cred*hist*length	Credit history length