

## **ABSTRACT**

Yuliana Vina Humira:

Final Assignment Report:

Model Refinement of a Credit Scorecard for Small and Medium Enterprises Loans of Bank X Using Logistic Regression

Risk management plays fundamental roles on bank activities, include credit risk, market risk, liquidity risk, and operational risk. New Frontier Solutions Pte. Ltd in co-operation with Bank X has developed a credit scoring model as credit risk management application. The credit scorecard considered three aspects, which are moral risk, business risk, and financial risk factors. The current credit risk model is measured its performance by calculating the GINI Coefficient and the Kolmogorov-Smirnov value also the model parameters are re-estimated. The new model is considered to be built not only by statistical approach which is using Logistic Regression method but applied the business logic as well. The model is validated by measuring the power discriminatory of the model to separate between “good” customer and “bad” customer and also carried out the population stability test of the model.

Keyword:

Credit Risk Management, Credit Scoring, Logistic Regression, GINI Coefficient, Kolmogorov-Smirnov.

## TABLE OF CONTENTS

COVER PAGE.....	i
APPROVAL SHEET .....	ii
STATEMENT OF AGREEMENT OF SCIENTIFIC WRITING PUBLICATION FOR ACADEMIC PURPOSES .....	iii
PREFACE.....	iiv
ABSTRACT.....	vi
TABLE OF CONTENTS.....	vii
LIST OF TABLES .....	ix
LIST OF FIGURES .....	xii
1. INTRODUCTION .....	1
1.1. Background .....	1
1.2. Research Question.....	3
1.3. Purpose Statement.....	4
1.4. Scope .....	4
1.5. Outline.....	4
2. LITERATURE REVIEW .....	6
2.1. Credit Scoring .....	6
2.2. Variable Selection .....	7
2.2.1. Contingency Table (Bhattacharya & Johnson, 1977).....	7
2.2.2. Pearson’s Chi Square Test (Bhattacharya & Johnson, 1977).....	9
2.3. Modeling The Probability of Default.....	9
2.3.1. Logistic Regression (Hosmer & Lemeshow, 2000) .....	10
2.3.1.1. Fitting The Multiple Logistic Regression Model (Hosmer & Lemeshow, 2000) .....	10
2.3.1.2. Testing for The Significance of The Model (Hosmer & Lemeshow, 2000).....	11
2.3.2. Analysis Using R (Everitt & Hothorn, 2010) .....	12
2.4. Model Validation.....	13
2.4.1. Power Discrimination .....	13
2.4.1.1. GINI Coefficient.....	13
2.4.1.2. Kolmogorov-Smirnov Test.....	14
2.4.2. Population Stability .....	14
3. METHODOLOGY .....	16
3.1. Variable Definition.....	16

3.2.	Data Set .....	17
3.3.	Measuring The Performance of Model 1.1 .....	18
3.4.	Modeling The Probability of Default .....	18
3.4.1.	Data Checking and Cleaning .....	21
3.4.2.	Univariable Analysis .....	22
3.4.3.	Pearson's Chi-square Test .....	22
3.4.4.	Multivariable Analysis.....	23
3.4.5.	Model Development .....	23
3.4.5.1.	Preliminary Main Effect Model.....	23
3.4.5.2.	Preliminary Final Model.....	24
3.4.6.	Building Scorecard .....	24
3.4.7.	Model Validation .....	25
3.4.7.1.	Power Discrimination .....	25
3.4.7.2.	Population Stability .....	26
3.5.	Conclusion.....	28
4.	DATA PROCESSING AND ANALYSIS .....	29
4.1.	Company Profile .....	29
4.2.	Credit Risk Model 1.1 .....	29
4.3.	Selection Variable .....	41
4.4.	Analysis of Parameter Estimation using Logistic Regression with R .....	44
4.5.	Credit Risk Model 1.2 .....	56
4.6.	Score Distribution .....	59
4.7.	Cut-off Rate.....	61
4.8.	Model Validation.....	63
4.8.1.	Discriminatory Power .....	64
4.8.1.1.	GINI Coefficient.....	64
4.8.1.2.	Kolmogorov-Smirnov test .....	66
4.8.2.	Stability Population .....	68
4.8.2.1.	Kolmogorov-Smirnov Goodness-of-fit .....	68
4.8.2.2.	Chi-square Goodness-of-fit .....	70
4.8.2.3.	$\chi^2$ Goodness-of-fit .....	71
5.	CONCLUSION AND REMARKS .....	73
5.1.	Conclusion.....	73
5.2.	Remarks.....	77
	REFERENCES .....	78

## LIST OF TABLES

1.1. Contribution Score for Each Factors of Credit Risk Model 1.1 .....	2
2.1. An $r \times c$ Contingency Table .....	8
2.2. Cell Probabilities .....	8
3.1. Data Set Model 1.1 .....	17
3.2. Data Set Model 1.2 .....	17
3.3. Calculation of GINI Coefficient .....	19
3.4. Rules of Quality Assessment by GINI Coefficient.....	19
3.5. Calculation of Kolmogorov-Smirnov Value .....	20
3.6. Rules of Quality Assessment by Kolmogorov-Smirnov .....	21
3.7. Calculation of Kolmogorv-Smirnov for Testing the Stability of Population.....	26
3.8. Calculation of Chi-square Goodness-of-fit for Testing the Stability of Population.....	27
4.1. Weight Contribution of Each Factor from Model 1.1 .....	30
4.2. Set Cut-off Rates for Model 1.1 .....	31
4.3. The Score Distribution of Default and Non-Default for Model 1.1 .....	32
4.4. Calculation of GINI Coefficient for Model 1.1 .....	34
4.5. Scorecard Quality Guidelines by GINI Coefficient.....	36
4.6. The Calculation of Kolmogorov-Smirnov Test Statistic for Model 1.1.	37
4.7. Scorecard Quality Guidelines by Kolmogorov-Smirnov Test Statistic..	38
4.8. Parameter Re-estimation of Model 1.1 .....	39
4.9. The Interpretation for Model 1.1 .....	40
4.10. Chi-square Calculation Summary.....	42
4.11. The Parameter Estimation of Mod.....	44

4.12. The Analysis of Deviance Mod1 .....	46
4.13. The Analysis of Deviance Mod2 .....	47
4.14. The Parameter Estimation of Mod1 .....	47
4.15. Interpretation of Positive Parameters for Mod1 .....	49
4.16. The Analysis of Deviance Mod3 .....	50
4.17. The Parameter Estimation of Mod3.....	50
4.18. The Analysis of Deviance Mod4 .....	51
4.19. The Parameter Estimation of Mod4.....	52
4.20. The Analysis of Deviance Mod5 .....	53
4.21. The Parameter Estimation of Mod5.....	53
4.22. The Analysis of Deviance Mod6 .....	54
4.23. The Parameter Estimation of Mod6.....	54
4.24. The Parameter Estimation of Mod7.....	55
4.25. Credit Risk Model 1.2 .....	57
4.26. Contribution Credit Score per Group Risk Factor .....	58
4.27. The Score Distribution of Credit Risk Modelk 1.2 .....	59
4.28. The Cut-off Rates of Model 1.2.....	61
4.29. The Interpretation of Credit Score Distribution.....	62
4.30. The Interpretation of Credit Score Distribution with Cut-off Score of 25 for Reject and 66 for Approval.....	63
4.31. The Interpretation of Credit Score Distribution with Cut-off Score of 30 for Reject and 66 for Approval.....	63
4.32. The Calculation of GINI Coefficient for Model 1.2.....	64
4.33. The Calculation of Kolmogorov-Smirnov Statistic for Model 1.2.....	66

4.34. The Calculation of Kolmogorov-Smirnov Goodness-of-fit for Model 1.2 .....	69
4.35. The Calculation of Chi-squared Goodness-of-fit for Model 1.2 .....	70
4.36. The Calculation of Chi-squared Goodness-of-fit (Adjusted) for Model 1.2 .....	71
5.1. The Parameter Re-estimation of Model 1.1.....	74
5.2. The Comparison Model 1.1 and Model 1.2.....	75

## LIST OF FIGURES

4.1. Distribution of Percentage Non Default and Default per Score Band for Model 1.1 .....	33
4.2. GINI Coefficient Graph of Model 1.1 .....	35
4.3. The Kolmogorov-Smirnov Graph of Model 1.1 .....	38
4.4. Distribution of Percentage Non Default and Default per Score Band for Model 1.2 .....	61
4.5. GINI Coefficient Graph of Model 1.2 .....	65
4.6. Kolmogorov-Smirnov Graph of Model 1.2 .....	67
5.1. The Distribution of Percentage Non Default and Default per Score Band for Model 1.1 .....	73
5.2. Score Distribution Credit Risk Model 1.2 .....	77

