

**COMBINING LOAN FRAUD DETECTION AND DEPOSIT RISK PROFILES TO  
IMPROVE THE EFFECTIVENESS OF DEPOSIT RECONCILIATION AND  
VERIFICATION PROCESS IN IDIC**

By

Arsandi Akhmad  
2-1751-008

MASTER'S DEGREE  
in

Master of Information Technology  
Faculty of Engineering and Information Technology



SWISS GERMAN UNIVERSITY  
The Prominence Tower  
Jalan Jalur Sutera Barat No. 15, Alam Sutera  
Tangerang, Banten 15143 - Indonesia

February 2019

Revision after the Thesis Defense on 24<sup>th</sup> January 2019

## STATEMENT BY THE AUTHOR

I hereby declare that this submission is my work, and to the best of my knowledge, it contains no material previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any other degree or diploma at any educational institution, except where due acknowledgment is made in the thesis.

Arsandi Akhmad

Student

Date

Approved by:

Dr. Ir. Lukas, MAI, CISA, IPM.

Thesis Advisor

Date

Dr. Eng Bagus Mahawan , B.Eng., M.Eng

Thesis Co-Advisor

Date

Dr. Maulahikmah Galinium, S.Kom.,M.Sc.

Dean

Date

Arsandi Akhmad

## ABSTRACT

### Combining Loan Fraud Detection and Deposit Risk Profiles To Improve The Effectiveness of Deposit Reconciliation And Verification Process By IDIC

By

Arsandi Akhmad

Dr. Ir. Lukas, MAI, CISA, IPM, Advisor

Dr. Eng Bagus Mahawan, B.Eng., M.Eng, Co-Advisor

SWISS GERMAN UNIVERSITY

IDIC as banking regulator in the mandate to handle the failed bank has two main tasks, namely completing all bank obligations or known as the liquidation of the bank and guaranteeing deposits that exist in the bank. In carrying out its duties as deposit guarantor, IDIC conducts a series of processes called reconciliation and verification. This study discusses how to increase the effectiveness of the reconciliation process and the verification of deposits doing by IDIC. Usually, this process is complicated because most banks that fail and liquidated because fraud and the most significant portion of fraud in the failed bank is loan fraud. The existence of fraud makes the validity of the balance recorded on the core banking application doubtful. It was often found in the flow of customer deposits that funds originate from loan fraud and to ensure that it takes a long time while the period of the reconciliation and verification process is limited to ninety working days.

For this reason, IDIC need a method that can increase effectiveness in this process, especially related to the detection of fraud. So in this study, a method was developed by using depositors' risk profiles combined with the detection of loan fraud that aims to increase the effectiveness of reconciliation and verification process in order to insured deposit.

In developing this method, data is needed following the process of reconciliation and verification of deposits as noted as in Standard Operation Procedure such as data related

to savings, deposits, loans and detailed data related to transactions. These data are processed and make the dataset in this study process and will use in loan fraud detection and depositors' risk profile.

This application uses the CRISP-DM framework with the Random Forest algorithm as the method. In this study also uses the SMOTE algorithm as handling data imbalance. The results of the analysis show that the features extracted or selected from failed bank data can be used in predicting loan fraud and insured deposit predictions with high accuracy and precision.

*Keywords: data analytics, machine learning, predictions, banks, fraud, loan fraud, dataset, reconciliation, verification, IDIC, deposit, validation, training set, dataset.*





© Copyright 2019

by Arsandi Akhmad

All rights reserved

SWISS GERMAN UNIVERSITY

## **DEDICATION**

I dedicate this research for my lovely wife (Marwanti) who always supports me every single time, for my beautiful daughters (Afiqah and Adifa) who always make me happy and accompanied me everywhere to do this research.



## ACKNOWLEDGEMENTS

I wish to thank the members of my committee for their support, patience and good humor. Their gentle but firm direction has been most appreciated. Greatest appreciation to Dr. Lukas, S.T., MAI. (Advisor) for the continuous support of my thesis and research in Master and Dr. Eng Bagus Mahawan, B.Eng., M.Eng (Co-Advisor), for his patience, motivation, enthusiasm, and immense knowledge. Their guidance helped me all the time of my research and writing of this thesis.

I have found my coursework throughout the Curriculum and Instruction program to be stimulating and thoughtful, providing me with the tools with which to explore both past and present ideas and issues.



SWISS GERMAN UNIVERSITY

## TABLE OF CONTENTS

	Page
STATEMENT BY THE AUTHOR.....	2
ABSTRACT.....	3
DEDICATION.....	6
ACKNOWLEDGEMENTS.....	7
TABLE OF CONTENTS.....	8
LIST OF FIGURES.....	10
LIST OF TABLES.....	11
CHAPTER 1 - INTRODUCTION.....	12
1.1 Background.....	12
1.2 Research Problems.....	14
1.3 Research Objectives.....	14
1.4 Research Questions.....	14
1.5 Hypothesis.....	15
1.6 Research Scope and Limitation.....	15
1.7 Significant of Studies.....	16
1.8 Thesis Structure.....	16
CHAPTER 2 - LITERATURE REVIEW.....	18
2.1 Indonesia Deposit Insurance Corporation.....	18
2.1.1 The Function and Role of the Deposit Insurance Corporation.....	19
2.1.2 Reconciliation and Verification Process.....	19
2.2 Liquidated Bank and Deposit.....	21
2.2.1 Customer Deposit.....	22
2.2.2 Deposit Risk Profile.....	22
2.3 Fraud in Financial Industry.....	23
2.3.1 Classification of Fraud.....	24
2.3.2 Fraud Triangle.....	25
2.4 Random Forest.....	26
2.4.1 Random Forest Regression Analysis.....	26
2.4.2 Random Forest Definition.....	28
2.4.3 Classification and Regression Trees (CART).....	29
2.4.4 Random Forest Algorithm.....	30



2.5	Machine Learning Dataset.....	32
2.5.1.	Feature Engineering.....	33
2.5.2.	Imbalance Data.....	33
2.5.3.	Data Normalization.....	34
2.5.4.	SMOTE.....	36
2.6	Previous Studies.....	37
CHAPTER 3 – RESEARCH METHODS.....		41
3.1	Materials and Equipment.....	41
3.2	Analytical Method.....	44
3.2.1	Business understanding.....	45
3.2.2	Data understanding.....	46
3.1.3	Data preparation.....	47
3.1.4	Modeling.....	48
3.1.5	Evaluation.....	49
3.1.6	Validation.....	50
CHAPTER 4 – RESULTS AND DISCUSSIONS.....		52
4.1	Business and Data Understanding.....	52
4.2	Data Preparation.....	54
4.2.1	Creating Dataset.....	55
4.2.2	Dataset Balancing.....	57
4.3	Data Analysis.....	58
4.3.1	Growing the Number of Trees.....	58
4.3.2	Selection of m values (Random Subset) based on the value of the error rate ..	60
4.3.3	Importance Variable.....	61
4.3.4	Random Forest Prediction Results.....	63
4.4	The Effect Of Loan Fraud Predictions On Deposit Insurance.....	64
4.4.1	Insured Deposit Prediction.....	64
CHAPTER 5 – CONCLUSIONS AND RECCOMENDATIONS.....		67
5.1	Conclusions.....	67
5.2	Recommendations.....	69
GLOSSARY.....		70
REFERENCES.....		72
APPENDIX A – R Code Loan Fraud Detection.....		74
APPENDIX B – R Code Deposit Risk.....		81
CURRICULUM VITAE.....		85