

**MODELING BANKING STABILITY INDEX USING MACHINE LEARNING
TECHNIQUE**

By

Agus Afiantara
21751009

MASTER'S DEGREE
in

MASTER OF INFORMATION TECHNOLOGY
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 January 2019

STATEMENT BY THE AUTHOR



THESIS CO-SUPERVISOR

Date

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

Dean

Date

Agus Afiantara

ABSTRACT



calculation need in the model. A very important point of concern in this research is to reduce dimension of variables selected using PCA as generally used to compress the data into a new dataset with fewer dimension. Bagged decisions trees like Random Forest is used to estimate the importance of features. Thorough the research Random Forest as machine learning technique and combined with PCA is shown to have the best result to classify and select the core variables of economic indicators and predict the

instability of financial system in Indonesia. Thirty three variables are used to predict the instability of financial system signal that constructed from Financial Stability Index (FSI) of Central Bank of Indonesia with monthly data. The algorithm is trained over the data from 2004-2011 period. As the result nine most components analysis are obtained as input for random forest machine learning to predict the instability of financial system





DEDICATION

I dedicate my thesis work to my God Allah SWT, and to my family, and many thanks to my lovely wife Linda Nirmala, who always support and care of me, and my daughter who I loved, Putri Zahra Maharani, and Putri Azizah Dafianti. I give you an example



ACKNOWLEDGEMENTS

I wish to thank to my fellows and my colleagues who want to be used as a place to discuss a lot about my thesis, there are my Bro Hendra Syamsir and Mas Advis Budiman always share the knowledge in helping me to finish this research. Thank you



TABLE OF CONTENTS

	Page
STATEMENT BY THE AUTHOR.....	2
ABSTRACT.....	3
DEDICATION.....	6



2.5.1. Random Forest Classifier.....	36
2.5.2. Random Forest Regressor.....	37
2.6. Related Works.....	39
CHAPTER 3 – RESEARCH METHODS.....	46
3.1. CRISP-DM Methodology.....	46
3.1.1. Business Understanding.....	48

3.1.2. Data Understanding	50
3.2. Data Preparation	51
3.2.1. NA Treatment	51
3.2.2. Log Transformation	51
3.2.3. Normalization	52



4.6.2. Hold-out validation result	70
4.6.3. Parameter Tuning	78
4.7. Final Model and Performance of BSI	84
4.8. Interpret Model and Discussion	88
4.8.1. Visualizing a Single Decision Tree	89
4.8.2. Variable Importance	89
4.8.3. The Estimated Model Based on K-Fold Cross Validation	91

4.9.	Save the model for future use	93
4.10.	Score New Data	93
4.11.	Lost Interpretation of PCA	93
CHAPTER 5 – CONCLUSIONS AND RECCOMENDATIONS		102
5.1.	Conclusions.....	102

