



TABLE OF CONTENTS

APPROVAL PAGE	i
PLAGIARISM STATEMENT.....	ii
FOREWORD	iv
TABLE OF CONTENTS.....	vi
LIST OF FIGURES	viii
LIST OF TABLES	ix
ABSTRACT.....	x
CHAPTER I INTRODUCTION	1
1.1 Research Background	1
1.2 Research Problem	2
1.3 Research Scope	3
1.4 Research Objectives.....	3
1.5 Research Advantages	3
CHAPTER II LITERATURE REVIEW	4
CHAPTER III THEORETICAL BASIS.....	9
3.1 Mobile Money.....	9
3.2 Fraud.....	10
3.3 Preprocessing.....	11
3.3.1 Data Transformation	11
3.3.2 Oversampling using SMOTE.....	12
3.3.2.1 Imbalance Data	12
3.3.2.2 Handling Imbalance Data.....	12
3.3.2.3 Synthetic Minority Over-sampling Technique.....	13
3.4 Support Vector Machine	15
3.5 Bayesian Belief Network	18
3.6 Decision Tree.....	19
3.7 Performance Evaluation.....	20
CHAPTER IV ANALYSIS AND DESIGN	22
4.1 General Description	22
4.2 Research Stages	22
4.2.1 Data Collection	22



4.2.2	Preprocessing	23
4.2.3	Split Test and Training.....	25
4.2.4	Modelling.....	25
4.2.5	Evaluation	26
CHAPTER V IMPLEMENTATION.....		28
5.1	Hardware and Software Specification.....	28
5.2	Implementation of Data Collection.....	28
5.3	Implementation of Preprocessing.....	28
5.4	Implementation of Splitting training and test data.....	30
5.5	Implementation of Support Vector Machine.....	31
5.6	Implementation of Gaussian Naïve Bayes	31
5.7	Implementation of Decision Tree using CART Algorithm.....	32
5.8	Implementation of the Evaluation.....	32
CHAPTER VI RESULT AND DISCUSSION		34
6.1	Preprocessing Result.....	34
6.2	Modelling Result.....	35
CHAPTER VII CONCLUSIONS		44
7.1	Conclusions.....	44
7.2	Recommendations.....	45
REFERENCES.....		46
APPENDIX A		49
APPENDIX B		54
APPENDIX C		59



LIST OF FIGURES

Figure 3. 1 Result of converting categorical attribute into binary	11
Figure 3. 2 Pseudo-Code of SMOTE	13
Figure 3. 3 Illustration of SMOTE process before synthetic data is added	14
Figure 3. 4 Illustration of SMOTE process after synthetic data is added	15
Figure 3. 5 SVM Illustration (Gao et al., 2009)	16
Figure 3. 6 Illustration of different Gamma values (Gao et al., 2009).....	17
Figure 3. 7 Example of Simple BBN Graph	18
Figure 3. 8 Example of Decision Tree Diagram (Witten and Frank., 2011).....	19
Figure 4. 1 Sample of Mobile Money Dataset	23
Figure 4. 2 Illustration Diagram of preprocessing	25
Figure 4. 3 Overall Workflow Diagram of the Research	27
Figure 5. 1 Implementation of standardization	29
Figure 5. 2 Implementation of converting categorical feature into biner	29
Figure 5. 3 Implementation of Oversampling using SMOTE.....	30
Figure 5. 4 Implementation of Split training and test data.....	31
Figure 5. 5 Implementation of Support Vector Machine Classifier.....	31
Figure 5. 6 Implementation of Gaussian Naïve Bayes Classifier	32
Figure 5. 7 Implementation of Decision tree using CART	32
Figure 5. 8 Implementation of confusion matrix and classification report calculation	33
Figure 5. 9 Implementation of Accuracy calculation.....	33
Figure 6. 1 Cash_IN and Transfer distribution diagram	35
Figure 6. 2 Distribution class from the original data	36
Figure 6. 3 Distribution class of testing data	38



LIST OF TABLES

Tabel 2. 1 Literature Study.....	7
Tabel 3. 1 Example of generation of synthetic examples (Chawla et al, 2002)....	14
Tabel 3. 2 Kernel and its function.....	16
Tabel 3. 3 Illustration of Confusion Matrix Table	20
Tabel 4. 1 Description of each attribute in dataset.....	22
Tabel 6. 1 Attributes Before Standardization.....	34
Tabel 6. 2 Attributes After Standardization	34
Tabel 6. 3 Confusion Matrix produce by each model.....	36
Tabel 6. 4 Distribution class table for each train dataset	37
Tabel 6. 5 Decision Tree before SMOTE	38
Tabel 6. 6 Decision Tree after SMOTE is applied.....	39
Tabel 6. 7 Support Vector Machine before SMOTE	40
Tabel 6. 8 Support Vector Machine after SMOTE	41
Tabel 6. 9 GaussianNB before SMOTE.....	42
Tabel 6. 10 GaussianNB after SMOTE.....	43
Table A. 1 Confusion Matrix Table of Decision Tree Before SMOTE is applied	49
Table A. 2 Confusion Matrix Table of Decision Tree After SMOTE is applied..	51
Table B. 1 Confusion Matrix Table of SVM Before SMOTE is applied	54
Table B. 2 Confusion Matrix Table of SVM After SMOTE is applied.....	56
Table C. 1 Confusion Matrix Table of GaussianNB Before SMOTE is applied..	59
Table C. 2 Confusion Matrix Table of GaussianNB After SMOTE is applied	60