TABLE OF CONTENTS

COVER

ORIGINALITY STATEMENT OF THE THESIS

THESIS SUPERVISOR'S APPROVAL

THESIS EXAMINATION COMMITTEE

ABSTRACT ..... **v**

ACKNOWLEDGEMENTS ..... **vii**

TABLE OF CONTENTS ..... **ix**

LIST OF FIGURES ..... **xi**

LIST OF TABLES ..... **xii**

CHAPTER I : INTRODUCTION **1**

1.1 Background 1

1.2 Identification of Problem 3

1.3 Research Scope 4

1.4 Problem Formulation 4

1.5 Research Objectives 5

1.6 Thesis Outline 5

CHAPTER II : LITERATURE REVIEW **7**

2.1 Fraudulent Financial Statements Detection 7

2.2 Fraudulent Financial Statements Predictors 9

2.3 Artitifical Neural Networks (ANN) 10

2.4 Synthetic Minority Over-Sampling (SMOTE) 12

2.5 Performance Evaluation Methods 12

CHAPTER III : RESEARCH METHODOLOGY **15**

3.1 Research Workflow 15

3.2 Data and Variables 16

3.3 Data Preprocessing 18

3.3.1 Serial Fraud 19

3.3.2 Minority Oversampling 19

3.4 Proposed Model 20

3.5 Experiments 21

3.6 Evaluation Metrics 22

CHAPTER IV : RESULTS AND DISCUSSION **23**

4.1 Correlation Analysis 23

4.2 Experiment Results 24
4.2.1 First experiment 27
4.2.2 Second experiment 28
4.2.3 Third experiment 30
4.2.4 Fourth experiment 33
4.2.5 Run-time Performance 34
4.3 Publication 35
CHAPTER V : CONCLUSION 45
5.1 Conclusion 45
5.2 Future Research 46
REFERENCES
APPENDIX A : CORRELATION MATRIX HEATMAP

LIST OF FIGURES

Figure 3.1 Research Workflow Diagram 15
Figure 4.1 Architecture of the Proposed Networks..... 27
Figure 4.2 Compilation of AUC Scores Over the Years..... 32

LIST OF TABLES

Table 3.1 Description of variables used in Cecchini approach 17
Table 3.2 Description of variables used in Dechow approach 18
Table 4.1 AUC score based on different numbers of neurons 28
Table 4.2 Summary of comparison with previous study (Test period 2003 – 2008) 30
Table 4.3 Summary of comparison with previous study (Test period 2003 – 2005) 31
Table 4.4 Summary of comparison with previous study (Test period 2003 – 2011) 31
Table 4.5 Summary of comparison with previous study (Test period 2003 – 2014) 31
Table 4.6 Ignoring serial fraud (Test period 2003 – 2008) 33
Table 4.7 Run-time Summary 34