FINANCIAL FRAUD DETECTION USING MACHINE LEARNING

Thilini Upeksha Kodituwakku

(148225U)

Degree of Master of Science in Computer Science

Department of Computer Science and Engineering

University of Moratuwa
Sri Lanka

March 2018

FINANCIAL FRAUD DETECTION USING MACHINE LEARNING

Thilini Upeksha Kodituwakku

(148225U)

Thesis submitted in partial fulfillment of the requirements for the degree Master of Science in Computer Science

Department of Computer Science and Engineering

University of Moratuwa
Sri Lanka

March 2018

Declaration

I hereby declare that this is my own work and this thesis/dissertation does not incorporate any material previously submitted for a Degree or Diploma in any other University or institute of higher learning without proper acknowledgement and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my thesis/dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future works (such as articles or books).

Signature: Date:

The above candidate has carried out research for the Masters/MPhil/PhD thesis/ Dissertation under my supervision.

Signature of the supervisor: Date:

Prof. Amal Shehan Perera

ABSTRACT

The method of performing transactions by means of payment cards is extremely efficient and the payment card industry is rapidly growing in popularity. However, the frauds associated with the payment cards are increasing and the patterns are evolving. Although a relatively smaller percentage is detected, fraud has become a major issue that affects the global banking industry. Machine learning techniques are widely used for payment card fraud detection.

The use of machine learning techniques generates successful results as there are large numbers of historical data that could be used for mining and manipulation. There are various machine learning algorithms available to construct fraud detection models. The main drawback of those models is their inability to deliver results accurately and efficiently at the level required by the industry as there is only a fine line between the fraudulent and non-fraudulent transactions.

The aim of this research is to create a model that reduces the present gap in the detection of payment card frauds using the ensemble machine learning technique. Ensemble methods are learning models that achieve performance by combining the opinions of multiple weaker models. The performance evaluation of a new ensemble model has been done on the real world financial data and the results indicated its capability of identifying a high percentage of frauds with low false alarm rate than the existing models in the payment card industry. Finally, results are analyzed, interpreted and directions for further research are suggested.

Acknowledgement

I would like to express my sincere gratitude to my supervisor, Prof. Amal Shehan Perera for providing me with the knowledge, support, and guidance throughout the duration of this project. Furthermore, I would like to thank my parents for the continuous support they gave me during this period. Also, I like to thank my colleagues who helped me immensely to gain an in-depth knowledge of payment card industry.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Description
ANN	Artificial Neural Networks
ATM	Automated Teller Machine
API	Application Programming Interface
EMV	Europay, MasterCard, Visa Standard
FDM	Fraud Detection Model
NFC	Near Field Communication
PIN	Personal Identification Number
POS	Point-Of-Sale
RFID	Radio-Frequency Identification
SMS	Short Message Service
SVC	Stored Value Card (also known as Prepaid Cards)