

Informatics Institute of Technology

In Collaboration with

University of Westminster, UK

Credit Card Fraud Detection using Machine Learning

A dissertation by

K.W.A Thushini Maleesha Ranasinghe

Supervised by

Mr. Rathesan Sivagnanalingam

Submitted in partial fulfilment of the requirements for the

BSc (Hons) in Software Engineering degree

Department of Computing

May 2019

© The copyright for this project and all its associated products resides with

Informatics Institute of Technology

Abstract

Credit card is one of the most used transactional methods in the world and every year, fraudulent credit card transactions cause losses of billions of dollars. Designing efficient systems for fraud detection is key to reducing these losses and more and more systems rely on machine learning techniques to assist fraud investigators. However, the design of fraud detection systems is particularly challenging because of the non-stationary distribution of data, the highly unbalanced distribution of classes and the unavailability of labelled transactions due to confidentiality issues.

This thesis aims to develop an intelligent approach that will increase the performance of a credit card fraud detection system when the dataset is heavily imbalanced and the model is overfitted. Through an in-depth literature study and industry survey, different techniques and technologies that could be used for the development of this system were identified and an in-depth evaluation and testing was performed to identify the most suitable techniques and technologies for this project. Accuracy of the system was tested on the basis of values output by the fraud detection system. Using Random Forest Classifier makes the system perform well, produce acceptable results and is therefore justified.

Key words: Credit Card Fraud Detection, Machine Learning, Random Forest, Classification Algorithms