Informatics Institute of Technology
In collaboration with
University of Westminster, UK.

Fraud Detection Solution for Financial Transactions
With Machine Learning

A Dissertation By
Thushara Madushanka Amarasinghe

Supervised by
Mr. Achala Aponsu

Submitted in partial fulfilment of the requirements for the
BEng (Hons) Software Engineering Degree
Department of Computing

April 2018

ABSTRACT

Fraud has become a trillion-dollar industry today. Some finance companies have separate domain expert teams and data scientists are working on identifying fraudulent activities. Data Scientists often use complex statistical models to identify frauds. However, there are many disadvantages to this approach. Fraud detection is not real-time and therefore, in many cases fraudulent activities are identified only after the actual fraud has happened. These methods are prone to human errors. In addition, it requires expensive, highly skilled domain expert teams and data scientists. Nevertheless, the accuracy of manual fraud detection mythologies is low and it is very difficult to handle large volumes of data. More often, it requires time-consuming investigations into the other transactions related to the fraudulent activity in order to identify fraudulent activity patterns. Finance companies are not getting adequate return of interest (ROI) despite the resources and money spent on these traditional methods. Most of the traditional fraud detection models focused on discrete data points. (User accounts, IP addresses devices, etc…) However, these methods are no longer sufficient for today's needs. As fraudsters and hackers are using more advance and cutting edge techniques to mask their fraudulent activities even from the sharpest eyes. These methods can only detect known types of attacks therefore an analytical approach is required to address these drawbacks of the traditional methods.

In order to address the above-mentioned issues in fraud detection domain, a new fraud detection system was introduced which uses an Artificial Neural Network to identify fraudulent transactions.

Data preprocessing has been carried out on the kaggle dataset available on https://www.kaggle.com/ntnu-testimon/paysim1 The ANN was built with three layers, one input layer, a hidden layer and an output layer. Keras library was used to build the ANN. Once the network is properly trained, the trained model is stored in a persistent storage. Later the same model has been used to build the ANN to detect frauds. ANN was able to predict with 90% accuracy. A REST API has been exposed with the use of Flask framework to submit transaction events. Implemented system was tested thoroughly with test data sets. final system was evaluated by domain experts. Design, implementation and documentation have been carried out in an effective and in an efficient manner.