A logo for college computing

Description automatically generated

**Assessment Cover Page**

|  |  |
| --- | --- |
| *Student Full Name* | Alexandra Mavrofoti |
| *Student Number* | Sbs22052 |
| *Module Title* | Strategic Thinking (HDip in Data Analytics - Feb 2024 - HCI cohort) |
| *Assessment Title* | Capstone Project Proposal |
| *Assessment Due Date* | 29th of March 2024 |
| *Date of Submission* |  |

**Declaration**

By submitting this assessment, I confirm that I have read the CCT policy on academic misconduct and understand the implications of submitting work that is not my own or does not appropriately reference material taken from a third party or other source.

I declare it to be my own work and that all material from third parties has been appropriately referenced.

I further confirm that this work has not previously been submitted for assessment by myself or someone else in CCT College Dublin or any other higher education institution.

Table of Contents

Contents

[Introduction 2](#_Toc158384946)

[Scope 2](#_Toc158384947)

[Objectives 2](#_Toc158384948)

[Ethical Consideration 4](#_Toc158384949)

[Boundaries and limitation 4](#_Toc158384949)

[Conclusion 5](#_Toc158384949)

[References 6](#_Toc158384950)

Introduction

Financial fraud poses a significant threat to the stability and integrity of financial systems worldwide. It is also a pervasive problem that affects businesses, consumers, and financial institutions alike. Fraudulent activities such as identity theft, credit card fraud, money laundering, and insider trading can result in substantial financial losses and damage to reputations. Detecting and preventing these fraudulent activities is paramount to maintaining trust and confidence in financial systems.

Scope

The scope of fraud analysis and detection is broad and encompasses various activities aimed at identifying, preventing, and mitigating fraudulent activities across different sectors and industries. Here are some key aspects of the scope of fraud analysis and detection:

Validating financial transactions and data analysis techniques in combination with forensic tools used to investigate suspected fraud incidents, uncover evidence, and identify perpetrators. This may involve analyzing evidence, such as emails, documents, and computer logs, to trace the origins of activities.

Fraud Prevention Strategies: Developing and implementing fraud prevention strategies and controls to minimize the risk of fraud occurring. This may involve strengthening internal controls, enhancing security measures and implementing fraud detection systems and technologies. Employee training and awareness is a must in this case.

Moreover, Monitoring and Improvement: Continuously monitoring and evaluating fraud detection and prevention measures to identify areas for improvement and adapt to emerging fraud trends and threats. This includes conducting periodic risk assessments, performance reviews and benchmarking against industry best practices.

At last, fraud analysis and detection is dynamic and evolving, requiring a comprehensive and proactive approach to effectively address the landscape of fraud risks and threats. Collaboration, innovation, and a commitment to ethical principles are essential for combating fraud and safeguarding organizations and individuals against financial and of course reputational losses.

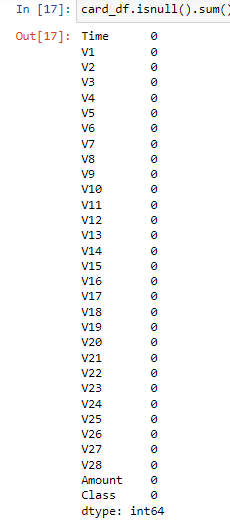
Objectives

Anomaly detection, also known as outlier detection, is a powerful technique used to identify patterns that deviate from normal behaviour within large datasets. In the context of financial fraud detection, anomaly detection algorithms analyze transactional data to detect unusual patterns or behaviours that may indicate fraudulent activity (unusual spending patterns, unauthorized access to accounts, or suspicious transactions etc.)

This capstone project will focus on developing a robust anomaly detection system for financial fraud prevention using a combination of machine learning algorithms and statistical techniques. The project will involve the following steps:

In the first semester, research is the first step where Data Collection and pre-processing is done: Gathering relevant financial transaction data from various sources and pre-processing it to ensure consistency and quality.

Data Analysis: Extracting informative features from the raw data to facilitate the detection of anomalies and data missing from our dataset. In this case there is some important information we can extract from the csv file (like how far was the transaction made from the holder’s home) and we need to pay special attention to. So we can apply different techniques for missing or invalid data (imputing missing values in our dataset, replacing NaN values with singe value etc.). However it looks like this is not needed here as there are no missing values.



Forecast using Model Selection and Training: Evaluating different anomaly detection algorithms like logistic regression (for supervised learning algorithms). This is commonly used for binary classification tasks in order to predict if a transaction is fraudlent or not.

For the second semester, report of findings and Performance Evaluation: Assessing the performance of the anomaly detection model using metrics. Continuously monitoring the performance of the deployed model and fine-tuning it to adapt to new fraud patterns and emerging threats. The summary of detected Fraud incidents must be detected during the reporting period (like types of fraud and impact).

Deployment and Integration: Integrating the trained model into a real-world financial system where the system, architecture and decision making mechanisms must be defined, the use of historical data and other parameters.

Delivery where the report and presentation is completed in order to do a final evaluation. For better understanding there will be analysis on the dataset downloaded from Kaggle website and some python scripts to get some statistics and graphs.

Ethical Considerations  
 Ethical considerations are paramount in fraud analysis and detection, as they guide professionals in maintaining integrity, fairness, and respect for privacy while carrying out their responsibilities. Here are some key ethical considerations in fraud analysis and detection:

Confidentiality: Fraud analysts must ensure the confidentiality of sensitive information obtained during investigations. They should only disclose information on a need-to-know basis and ensure that sensitive data is protected from unauthorized access.

Compliance with Laws and Regulations: Analysts must adhere to relevant laws, regulations, and industry standards governing fraud detection and investigation. This includes laws related to data privacy, evidence handling etc.

Transparency: The methods used for fraud detection should be transparent to those affected by them. Individuals should be aware of what data is being collected, how it's being used, and what algorithms or techniques are being employed to detect fraud.

Avoid Bias: Fraud detection algorithms should be designed and tested to minimize biases as they can lead to discriminatory outcomes, disproportionately affecting certain groups based on factors like race, gender or status.

Boundaries and limitations

Fraud analysis and detection, while crucial in mitigating financial losses and protecting against illicit activities, have certain boundaries and limitations that professionals must be mindful of. Here are some of them:

Legal Constraints on the methods and techniques that can be used in fraud analysis and detection. Analysts must adhere to laws and regulations governing privacy, data protection, and evidence handling, which may limit their ability to gather certain types of information or employ certain investigative tactics.

Resource Constraints which means limited resources, including time, budget, and manpower, can constrain the effectiveness of fraud analysis and detection efforts. Analysts may not have access to the necessary tools, technologies, or expertise needed to conduct comprehensive investigations.

Conclusion

In conclusion, fraud analysis is an essential component of risk management and security strategies, enabling organizations to identify, prevent, and mitigate fraudulent activities effectively. By investing in robust fraud detection and prevention measures, businesses and financial institutions can safeguard their assets, comply with regulations, reduce the cost and provide financial stability and maintain trust and confidence among stakeholders.

# 

# References

Machine Learning:

Mitchell, T.M. (1997). Machine Learning. McGraw-Hill Education.

Fraud Analysis:

Wells, J.T. (2016). Principles of Fraud Examination. John Wiley & Sons.

Doyle, J. T., & Feldman, R. (2016).Detecting Fraud in Organizations: Techniques, Tools, and Resources*. John Wiley & Sons.*

Kumar Abhishek , Dr Mounir Abdelaziz (2023). Machine Learning for Imbalanced Data: Tackle imbalanced datasets using machine learning and deep learning techniques

Repository:

https://github.com/Alexandra-Mavrofoti/Strategic-Thinking-CA\_1.git

Data Source:

https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud