SCHOOL OF COMPUTING
Faculty of Engineering

Project Proposal Form MCSD 6215
Sem: 1 Session: 2024/2025

SECTION A: Project Information.

Program Name:	**Masters of Science (Data Science)**
Subject Name:	**Project 1 (MCSD 6215)**
Student Name:	Nur Adriana Batrisyia binti Mohd Subri
Matric Number:	MCS241007
Student Email & Phone:	nuradrianabatrisyia@graduate.utm.my
Project Title:	Money Laundering Detection using Machine Learning Algorithm
Supervisor 1:	
Supervisor 2 / Industry Advisor(if any):	

SECTION B: Project Proposal

Introduction:

According to Bank Negara Malaysia, money laundering is a method to transform 'dirty' illegal money into 'clean' legitimate appearance. The money may come from criminal activities such as drug trafficking and corruption, thus the offenders need to conceal its unlawful origin before they can luxuriously spend the money. In general, the money laundering process can be summarized into three steps which are placement, layering and integration. Firstly, illegal money is placed into financial institution. Then, the money is transferred multiple times across many layers of accounts. Finally, after going through a cycle of complex transactions to disguise the origin, the money is then integrated back into the economy as lawful funds. The complexity of money laundering activities encourages the use of machine learning algorithm as a promising approach to identify various patterns and detect money laundering activities in large datasets.

Problem Background:

The United Nations Office on Drugs and Crime (UNDOC) estimated that around 2% to 5% of the world's total economic output is laundered globally every year. It brings to the significant amount of $800 billion - $2 trillion 'dirty' money. As of 2023, Malaysia received 317,435 Suspicious Transaction Report (STR) which was a 31% increase from last year where the reports are mainly on fraud, money laundering and tax offences. More than 100 individuals were arrested and RM290 million was recovered. Furthermore, 59,684 suspected mule accounts were identified and disrupted to hinder the process of disguising the origin of illicit funds vias multiple layering transactions. In order to effectively combat money laundering activities, this project seeks to explore more advanced approach using machine learning algorithm to learn complex transaction patterns and enhance money laundering detection.

Problem Statement:

Malaysia's Anti-Money Laundering and Counter Financing Terrorism (AML/CFT) regime is generally in compliance with Financial Action Task Force (FATF), however in terms of money laundering investigation and prosecution, it is still not showing a significant effectiveness even though the number of investigations is increasing. The total of money laundering prosecutions and convictions is still low, and Malaysia is not adequately targeting high-risk offences especially if it involves cross border transactions. This is because most financial institutions are using rule-based techniques to detect money laundering activities, but it is not powerful enough to identify the complex and hidden schemes used by criminals, especially in cross border transactions. Hence, there is a need to develop a machine learning approach to combat and stay ahead of sophisticated money laundering methods.

Aim of the Project:

This project aims to utilize supervised machine learning algorithm to efficiently detect money laundering activities as an effort to maintain financial integrity in Malaysia.

Objectives of the Project:

1) To perform data preprocessing and exploratory data analysis (EDA) to handle noisy data and understand data distributions.
2) To implement machine learning algorithms to learn patterns, identify anomaly transactions and detect money laundering activities.
3) To evaluate model using metrics such as True Positive Rate (TPR), False Positive Rate (FPR), True Negative Rate (TNR), False Negative Rate (FNR) and Area Under the Curve (AUC).

Scopes of the Project:

Since it is difficult to obtain real financial transaction data due to privacy reasons and legal constraints, hence this project is using synthetic money laundering datasets called SAML-D developed by Berkan Oztas and five other researchers in their paper entitled 'Enhancing Anti-Money Laundering: Development of a Synthetic Transaction Monitoring Dataset'. This dataset includes 28 typologies of transactions which is the type of money flow after it is first placed into the financial institution, thus brings greater realism to the dataset. Plus, it also has geographic locations which are important to analyze cross border transactions.

This project will use machine learning algorithms such as Support Vector Machine (SVM) or Random Forest to detect money laundering activities. First, data preprocessing and exploratory data analysis are performed to have a better understanding of the data and discover patterns and indicators for money laundering activities. Next, machine learning algorithm is applied to detect suspicious transactions and flag as money laundering activities. The algorithm will then be evaluated by using metrics such as True Positive Rate (TPR), False Positive Rate (FPR), True Negative Rate (TNR), False Negative Rate (FNR) and Area Under the Curve (AUC).

Expected Contribution of the Project:

1) Improve efficiency in detecting money laundering activities in complex situations and cross-border transactions.
2) Reduce false alarms by ensuring a low false positive rate compared to current rule-based approach.

Project Requirements:

Software: .....

Hardware: .....

Technology/Technique/ Methodology/Algorithm: .....

Type of Project (Focusing on Data Science):

[ ] Data Preparation and Modeling

[ ] Data Analysis and Visualization

[ ] Business Intelligence and Analytics

[ ] Machine Learning and Prediction

[ ] Data Science Application in Business Domain

Status of Project:

[ ] New

[ ] Continued

If continued, what is the previous title? .....

SECTION C: Declaration

I declare that this project is proposed by:

[/] Myself

[ ] Supervisor/Industry Advisor ( )

Student Name: .....

....

Signature **Date**

SECTION D: Supervisor Acknowledgement

The Supervisor(s) shall complete this section.

I/We agree to become the supervisor(s) for this student under aforesaid proposed title.

Name of Supervisor 1: .....

....

Signature **Date**

Name of Supervisor 2 (if any): .....

....

Signature **Date**

SECTION E: Evaluation Panel Approval

The Evaluator(s) shall complete this section.

Result:

[ ] FULL APPROVAL [ ] CONDITIONAL APPROVAL (Major)*

[ ] CONDITIONAL APPROVAL (Minor) [ ] FAIL*

***** Student has to submit new proposal form considering the evaluators' comments.

Comments:

Name of Evaluator 1:

Signature **Date**

Name of Evaluator 2:

Signature **Date**