ASIAN ACADEMY of MANAGEMENT JOURNAL of ACCOUNTING and FINANCE

AAMJAF Vol. 22, No. 1, 121–155, 2026

THE IMPACT OF KEY ECONOMIC INDICATORS ON MONEY LAUNDERING RISK IN MALAYSIA: A COMPARATIVE STUDY OF MACHINE LEARNING ALGORITHMS

Nurulhuda Ramli[1*], Heri Kuswanto[2] and Mohamad Sharimie Arif Mat Taib[3]

[1]Mathematics Section, School of Distance Education, Universiti Sains Malaysia, 11800 USM Pulau Pinang, Malaysia
[2]Department of Statistics, Institut Teknologi Sepuluh Nopember (ITS), 60111 Sukolilo, Surabaya, Indonesia
[3]AHAM Asset Management Berhad, Menara Boustead, 69 Jalan Raja Chulan, 50200 Kuala Lumpur, Malaysia

*Corresponding author: rnurulhuda@usm.my

Submitted: 31 Oktober 2025; Accepted: 27 January 2026; Published: 28 June 2026

To cite this article: Ramli, N., Kuswanto, H., & Mat Taib, M. S. A. (2026). The impact of key economic indicators on money laundering risk in Malaysia: A comparative study of machine learning algorithms. *Asian Academy of Management Journal of Accounting and Finance, 22*(1), 121–155. https://doi.org/10.21315/aamjaf2026.22.1.4

To link to this article: https://doi.org/10.21315/aamjaf2026.22.1.4

ABSTRACT

This study investigates the impact of key economic indicators on money laundering risk in Malaysia through a comparative analysis of machine learning (ML) algorithms. The study utilises three ML techniques, namely Random Forest (RF), Naïve Bayes (NB) and Logistic Regression (LR) to predict Malaysia's money laundering risks based on a dataset from 2000 to 2024. The dependent variable, money laundering risk, is analysed against a set of independent economic indicators, including Gross Domestic Product (GDP) per capita, exchange rate, annual inflation, Consumer Price Index (CPI), net secondary income, trade-related metrics, tax revenue and unemployment rate. To address class imbalance and enhance model robustness, data pre-processing, resampling, cross-validation and hyperparameter tuning were systematically implemented. Model performance was evaluated using multiple classification metrics to capture different aspects of predictive accuracy. Results demonstrate that NB consistently outperformed the other algorithms, achieving the strongest overall classification accuracy, agreement beyond chance, lowest prediction error and highest discriminatory power. Feature importance analysis using Recursive Feature Elimination (RFE) with LR algorithm identified the unemployment *rate, import volume index, and exchange rate as the most influential predictors of money laundering risk. These findings underscore the value of probabilistic ML approaches and feature selection techniques in enhancing anti-money laundering (AML) risk assessment and provide actionable insights for policymakers and regulatory authorities in Malaysia.*

Keywords: Economic indicators, Feature importance, Machine learning, Anti-money laundering, Financial crime

INTRODUCTION

Money laundering is the process of disguising the illicit origins of funds derived from criminal activities, such as drug trafficking, corruption, tax evasion and terrorism financing. The goal is to integrate these "dirty" funds into a clean and untraceable stream within the legitimate financial system. According to the United Nations Office on Drugs and Crime (UNODC), an estimated 2%–5% of global GDP, or between USD800 billion and USD2 trillion, is laundered worldwide annually (UNODC, n.d.). This vast scale demonstrates the significant threat of this activity to both financial stability and international security. The global challenge of money laundering carries severe impacts, not only economically but also socially and politically.

At its core, money laundering involves three primary stages; placement (placing illicit funds into the financial system), layering (cutting traces following complex transactions) and integration (re-integrating “cleaned” money into legitimate economy) (Reza et al., 2024; Schneider & Windischbauer, 2008). International organisations, such as the Financial Action Task Force (FATF), have developed standards to counter money laundering by fostering collaboration among financial and other sectors worldwide and encouraging the implementation of standardisation of anti-money laundering (AML) regulations. These efforts include procedures such as asset tracing, provisional measures for freezing illegal assets, and coordinated asset recovery (Jin & Wang, 2021). Countries that fail to implement or have poor AML system would create opportunities for criminals to exploit the system, thereby undermining global efforts to combat money laundering.

In relation to the South East Asian region, Malaysia has faced increasing scrutiny in terms of exposure to money laundering risks. According to the Basel AML Index, an independent annual ranking that assesses countries’ exposure to money laundering and terrorist financing (TF) worldwide, Malaysia has exhibited mixed trends since the index was first released in 2012. Most recently, Malaysia was ranked 66th out of 177 countries in the 2025 Basel AML Index. The country's money laundering risk score increased from 5.21 to 5.60 (on a scale of 0 to 10, where higher scores indicate greater risk) compared to the previous year. Notwithstanding this upward trend, Malaysia continues to be classified as a medium-to-high-risk jurisdiction for money laundering. This persistent vulnerability underscores Malaysia's ongoing exposure to money laundering risks at the international level, indicating that it is not yet considered a low-risk or "clean" jurisdiction in terms of AML performance.

In Malaysia, AML practices are governed by the Anti-Money Laundering, Anti-Terrorism Financing and Proceeds of Unlawful Activities Act 2001 (AMLATFPUAA, previously known as AMLA). This legislation was enacted in 2002, with subsequent amendments in 2007 and 2013, to combat money laundering, terrorism financing and proceeds for unlawful activities. Malaysian commercial banks are required to comply with strict regulations to prevent the generation of funds from illegal activities as well as to protect customers. They are also required to adhere to foreign exchange restrictions introduced in 1998, where all cross-border transactions exceeding RM5, 000.00 must be reported to the central bank (Mahmud & Ismail, 2023). As per Section 14 of AMLA, banks must promptly report to the central bank, Bank Negara Malaysia, about any transaction exceeding RM50,000 or any circumstance that raises suspicion of proceeds from unlawful activities.

Despite the stringent laws and regulations, the increasing complexity of financial crimes requires proactive and adaptive approaches for banks to mitigate money laundering risk. Critically, these efforts must look beyond internal transactional data to consider the macro level. Emerging studies highlight that money laundering risk is not solely a financial or legal issue but is also influenced by broader economic factors. Financial secrecy, auditing standards, political stability, and legal risk substantially determine a country's vulnerability to money laundering (Ghulam & Szalay, 2024). These indicators, when considered together with economic metrics like Gross Domestic Product (GDP) per capita, foreign trade and unemployment, provide a complex picture of systemic risk (Vagaská et al., 2022). Incorporating such variables into a sophisticated predictive mechanism can significantly enhance AML strategies, offering insights that are often unavailable from monitoring financial transactional or regulatory sources alone.

Recent research has demonstrated that machine learning (ML) techniques have the potential to improve current AML predictive models. ML is able to identify laundering patterns from large amounts of variables more effectively than traditional algorithms (Jullum et al., 2020). However, despite this growing body of knowledge, most existing ML research still mainly focuses on cross-country comparison (e.g., Ghulam & Szalay, 2024; Lokanan, 2024a) and places less emphasis on country-specific predictive model development, particularly with respect to Malaysia. In addition, it remains unclear how key economic indicators contribute to money laundering activity in Malaysia. The aim of this study is to fill this gap by examining the link between economic indicators and money laundering risk in Malaysia using ML algorithms.

LITERATURE BACKGROUND

Key Economic Indicators on Money Laundering

The AML framework have traditionally relied heavily on micro-level financial indicators, such as transaction monitoring, banking records and Know-Your-Customer (KYC) protocols to identify suspicious activities (Ghulam & Szalay, 2024). While these indicators are crucial for detecting illicit behaviour through abnormal transaction patterns, this narrow focus often yields high false positive rates, overwhelming compliance departments and diverting resources. Furthermore, sophisticated laundering operations increasingly exploit legal and regulated channels, including real estate and trade finance, which do not always exhibit clear transactional red flags (Masum et al., 2024; Tiwari et al., 2020). Beyond these micro-level anomalies, macroeconomic factors significantly influence the scale and complexity of money laundering, creating environments that either facilitate or deter illicit financial activities (Hendriyetty & Grewal, 2017; Tanzi, 1997).

Research has consistently shown that macroeconomic instability, weak institutions, and opaque financial systems create favourable conditions for financial crime, including money laundering (Schneider & Windischbauer, 2008; Walker & Unger, 2009). The impact of key macroeconomic performance indicators such as GDP on money laundering is a well-explored area in various studies, consistently revealing the significant scale and detrimental effects of illicit financial activities globally. Money laundering constitutes a substantial portion of global economic activity, with the International Monetary Fund (IMF) estimating that it accounts for between 2% and 5% of global GDP annually—equivalent to as much as USD2 trillion in illicit funds laundered worldwide (Nazar et al., 2024; Summe, 2000). This figure highlights the scale of the problem and its widespread negative impact on legitimate financial systems. Furthermore, the relationship between GDP and money laundering is shaped by vulnerabilities within the financial system, as the financial sector often serves as the main channel for laundering illicit funds, thereby posing substantial risks to global financial stability, particularly when such activities go undetected (Sarigul, 2013). Beyond GDP, other economic performance indicators also play a crucial role. Exchange rate volatility, for example, may incentivise individuals or companies to engage in currency conversions or cross-border transactions specifically designed to mask illicit activities (Avdjiev et al., 2019). This is because fluctuating exchange rates can make it easier to obscure the true value of transferred funds, creating opportunities for illicit parties to disguise the origin and movement of money.

Macroeconomic stability is another critical dimension influencing money laundering risk, with inflation and the Consumer Price Index (CPI) serving as particularly relevant indicators. High inflation destabilises an economy by rapidly eroding purchasing power and asset values, encouraging individuals to seek alternative ways to preserve their wealth, including moving it across borders or into less regulated channels. Such behaviour increases reliance on informal financial mechanisms and reduces transaction transparency, thereby elevating the risk of money laundering (Hendriyetty & Grewal, 2017). Several studies support this relationship. Shah and Aish (2022), using data from South Asian countries, identified a strong link between inflation, corruption and money laundering, showing that illicit financial activities can both respond to and intensify inflationary pressures. This demonstrates a bidirectional relationship where illicit financial activities can not only be influenced by inflation but can also contribute to it. Similarly, as a standard measure of inflation, the CPI directly influences monetary policy and financial systems, shaping the economic environment in which laundering activities occur. As inflation affects the purchasing power of money, it naturally impacts money laundering activities, therefore prompting launderers to adjust their strategies in response to shifting economic incentives (Khelil et al., 2024). Furthermore, evidence from

Shen and Dong (2019) indicates that CPI movements are significantly correlated with changes in the money supply, providing insights into how inflationary dynamics can create an environment conducive to money laundering activities.

Considering external sector indicators, net secondary income, such as remittances, can create significant vulnerabilities. Large volumes of cross-border transfers in the form of remittances, particularly in economies with underdeveloped regulatory frameworks, may be exploited by criminal networks to disguise illicit funds as legitimate flows (Hendriyetty & Grewal, 2017). The relationship between remittance flows and money laundering is further underscored by studies analysing the impact of transaction costs. For instance, Ahmed et al. (2021) showed that reducing remittance costs increase the volume of formal remittances, suggesting that easing these costs may help reduce reliance on informal channels that are often exploited for money laundering. Similarly, trade volumes, particularly exports and imports, are highly relevant indicators due to the prevalence of trade-based money laundering (TBML). Trade transactions are frequently exploited through fraudulent over- or under-invoicing, where the price, quantity or quality of goods is misrepresented to move value illicitly across borders, making it difficult to detect the true nature of the transactions (Ghulam & Szalay, 2024). This vulnerability is further compounded by the intersection with informal financial channels. Jayasekara (2023), in a study focusing on Sri Lanka, examined the impact of TBML and informal remittance services on the sustainability of the balance of payments. This research highlights how informal remittance services, often utilised due to regulatory shortcomings, contribute to TBML and pose significant challenges to economic stability. Further reinforcing this connection, Tran (2024) investigated the role of remittance costs in facilitating TBML by examining the relationship between high remittance costs and the over-invoicing of trade in Vietnam. The findings indicate that reducing remittance costs can help mitigate TBML by encouraging the use of formal financial channels.

Fiscal and social indicators provide further insights into money laundering risk. The unemployment rate is a significant factor, as elevated unemployment often signals widespread economic distress. This condition can drive individuals or groups toward illicit financial activities as a means of survival, thereby expanding the underground economy where laundered funds can be more easily hidden and integrated (FiroozAbadi et al., 2015).

While not directly addressing money laundering, studies like Fallahi et al. (2012) highlighted how unemployment and its volatility affected crime rates, suggesting a broader influence on criminal behaviour that can include money laundering. Phillips and Land (2012) found that higher economic crime rates often— influenced by unemployment—were associated with increased money laundering activities. Schneider and Windischbauer (2008) presented quantification techniques for measuring money laundering and discussed its social and economic impacts, including a significant link to economic factors like unemployment. Closely related in this context are the shadow economy and tax revenue. A large shadow economy, characterised by unreported economic activities and weak tax collection mechanisms, creates an ideal environment for money laundering. In such settings, illicit funds can be seamlessly integrated into the legitimate economy in the absence of effective oversight (Hendriyetty & Grewal, 2017; FiroozAbadi et al., 2015; Rider, 2019). This dynamic not only diminishes government resources but also weakens the rule of law. The interlinkage between tax evasion and money laundering is crucial; recognising tax evasion as a predicate offense for money laundering can significantly strengthen legal frameworks and enforcement efforts (Khan & Akhtar, 2022; Schlenther, 2013).

At the national level, money laundering harms a nation's economy through systemic consequences that extend far beyond the direct value of laundered funds. It can lead to capital misallocation by diverting resources to unproductive or criminal enterprises, resulting in inflated asset prices—particularly in real estate—eroded investor confidence, and substantial challenges for monetary policy management (Masum et al., 2024). Such illicit flows also tend to widen income inequality and create a long-term drag on productivity, as they reduce government revenue and distort economic competition (Hendriyetty & Grewal, 2017). These effects highlight the need to accurately assess how macroeconomic variables influence both the risk and scale of money laundering. To be effective, an AML framework must move beyond merely monitoring individual financial transactions and more fully account for broader macroeconomic conditions, structural imbalances, and institutional weaknesses that create an enabling e environment for financial crime (Rider, 2019). Integrating this macro-financial perspective is therefore crucial for developing predictive models that capture the overall systemic risk to national financial integrity.

Applications of Machine Learning (ML) in Money Laundering

Recent years have seen a surge in the use of analytical technologies to generate deeper insights for AML. Advanced AML technologies offer transformative potential, significantly enhancing capabilities in the identification and tracking of suspicious transactions, where traditional investigation methods would inevitably miss (Nazri et al., 2019). Technologies like AI and ML were foundational to this change, transforming the AML landscape from rigid, rule-based systems into ones that can intelligently and continuously "learn" to detect subtle anomalies. Their power lies in their ability to learn complex, non-linear relationships between various transaction features (like amount, frequency, geography and counterparty) and known outcomes (legitimate or suspicious). This allows them to precisely flag deviations from normal behaviour and effectively minimise false positives compared to simple rule-based systems (Jullum et al., 2020; Boateng et al., 2025; Domashova & Mikhailina, 2021).

Key ML approaches in AML are primarily categorised into supervised and unsupervised methods. Supervised learning can be divided into two main tasks: classification and regression, with classification being particularly relevant in AML contexts. Classification models are used to determine the group membership of input data, such as labelling transactions or customers as "suspicious" or "non-suspicious." These models were shown to be effective for finding complex behaviours as they are capable of learning patterns from labeled datasets.

Random Forest (RF) is an ensemble classification learning method that constructs a multitude of decision trees during training and outputs the mode of the predictions from the individual trees. By combining multiple "weak learners" into a single "strong learner," RF reduces the risk of overfitting and effectively captures non-linear relationships in the data. This capacity for capturing high-dimensional data and recognising subtle patterns has made RF highly effective in financial records. For example, Zhang et al. (2023) applied RF to establish a national AML index with an excellent 86.31% accuracy. RF is also used to forecast suspicious transactions in bank data, as demonstrated by studies from Lokanan (2019), Ketenci et al. (2021), and Huong et al. (2024). However, the model's success often depends on data preparation; Zhang and Trubey (2019) highlighted that the sampling plan is crucial for detecting rare events when using RF for money laundering. The versatility of RF extends into modern financial frontiers beyond traditional banking, it shows great effectiveness for cryptocurrency money laundering detection (Alotibi et al., 2022). Similarly, Mohan et al. (2023) combined RF with a convolutional network for identifying illegal cryptocurrency transactions and achieved better results than traditional approaches. Phyu and Uttama (2024) found that the RF model, when combined with a dataset balancing technique, produced the best results, achieving over 91% accuracy and 64% recall, particularly in capturing typological features associated with money laundering patterns.

While RF excels at complexity, Naïve Bayes (NB) offers a highly efficient alternative based on probabilistic logic. NB is a classification technique based on Bayes' Theorem, which assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature (the "independence" assumption). Despite its simplicity, it is remarkably effective for high-dimensional datasets. For instance, Tiwari et al. (2020) demonstrated NB's capability in distinguishing financial transactions as genuine or illegitimate. In addition, Lokanan (2019) found that NB, along with RF, was one of the best-performing algorithms for predicting money laundering transactions within bank systems. In specific scenarios, NB may even outperform its counterparts, as Liu and Chen (2024) showed NB to be superior to RF and Decision Tree (DT) in detecting intermediate accounts involved in money laundering. Furthermore, Gupta et al. (2022) demonstrated the effectiveness of NB in credit card fraud detection with an accuracy of around 80.4%. These findings reinforce the applicability of NB in financial fraud detection.

Logistic Regression (LR) remains a foundational tool to predict a binary outcome (e.g., "laundering" vs. "not laundering") by mapping predicted values to probabilities using a sigmoid function. Its mathematical transparency makes it highly valuable for regulatory and auditing purposes. A study conducted by Harris et al. (2021) showed the effectiveness of LR for predicting customers at risk of money laundering in a closed banking dataset of a Canadian bank. Kumar et al. (2021) also presented that LR, NB, k-Nearest Neighbour (k-NN) and DTs were the best models to detect laundering using both transactional-level and individual data. Moreover, Ghulam and Szalay (2024) used multinomial LR in investigating factors influencing AML risk. Lokanan (2024b) demonstrated that LR, together with SVM, is able to estimate the country's vulnerability to sanctions by encompassing a range of input variables such as financial transparency, political stability and unemployment rates.

Meanwhile unsupervised and semi-supervised ML methods are invaluable for anomaly detection and identifying money laundering patterns, particularly when the availability of labelled data is scarce or nonexistent. These techniques include clustering, isolation forest, and advanced graph-based models, which excel at discovering hidden structures and relationships within complex financial datasets (Yang et al., 2023; Labanca et al., 2022; Lokanan, 2024a; 2024b; Karim et al., 2024; Jensen & Iosifidis, 2022). In the context of AML, supervised ML is more widely adopted than unsupervised approaches. A primary advantage of supervised learning is that it aligns more effectively with compliance workflows and the rigorous requirements for explainability (Canhoto, 2020). Because these models learn explicit mappings from features to specific risk outcomes, the results can be linked to human-understandable factors that are easier to justify to regulatory bodies (Cardoso et al., 2022; Kute et al., 2021). Therefore, by utilising a supervised framework, this study can more effectively map the complex relationships between macroeconomic indicators and money laundering risk, ensuring that the resulting predictive model is both technically robust and transparent enough for policy application.

Macroeconomic Analysis of AML in the Malaysian Context

In the Malaysian context, empirical studies on AML have predominantly taken micro-level approaches, focusing on institutional structures and behavioural factors that influence money laundering risk. For example, Tasan et al. (2023) conducted a quantitative survey of 102 commercial bank employees to identify factors contributing to money laundering activities. Using descriptive and inferential statistical analysis, the study found that technology adoption, legal framework, income level and ethical behaviour were all positively correlated with laundering risk within the banking sector. Although the study focused on internal compliance behaviour, the inclusion of income level as a variable indirectly reflects macroeconomic dimensions such as income inequality and financial modernisation, suggesting that broader economic conditions may influence laundering tendencies in Malaysia's financial system. In contrast, Yusoff et al. (2023) examined high-profile cases such as 1Malaysia Development Berhad (1MDB), Shafee's case, and Genneva Malaysia Sdn. Bhd. through the lens of the fraud framework. Their study identified key elements such as pressure (often driven by financial distress or greed), opportunity (arising from weak economic controls or regulatory gaps), rationalisation and capability as critical drivers of laundering activity. This analysis demonstrates how economic incentives and systemic vulnerabilities interact, showing that weaknesses in financial governance and oversight can facilitate the concealment and circulation of illicit proceeds within Malaysia's economy.

Although direct studies linking macroeconomic variables to money laundering in Malaysia are limited, related financial research provides valuable insights. Sairin et al. (2020) investigated the impact of macroeconomic indicators on corporate financial distress using 10 years of panel data from Bursa Malaysia. Their econometric analysis found that money supply (M2), inflation (CPI), producer price index (PPI) and real interest rate (RIR) significantly influence financial vulnerability, with M2 and RIR showing the strongest effects. While not AML-specific, these findings provide a methodological foundation for integrating macroeconomic factors into AML risk models, as liquidity, inflation and interest rate movements are closely related to capital flows and the ease of fund concealment.

A notable advancement in predictive modeling and macro-financial analysis comes from Awang Abu Bakar (2024), who applied ML techniques, specifically RF and XGBoost models to assess the determinants of Central Bank Digital Currency (CBDC) adoption using data from the Bank for International Settlements (BIS) and the International Monetary Fund (IMF). The study identified cryptocurrency prevalence, cash circulation, and trade openness as critical predictors, achieving predictive accuracies of 83% and 80%, respectively. Although focused on financial innovation rather than AML, this research illustrates the potential of machine learning to process complex, macro-level datasets to predict national financial outcomes within the Malaysian context.

While these studies offer valuable insights into institutional and behavioural and financial systems, there remains a significant gap in empirical research that applies a data-driven, quantitative approach to directly link macroeconomic indicators to money laundering risk within the broader Malaysian context. Although macro-data have been successfully applied to topics such as digital currency adoption (Awang Abu Bakar, 2024), the influence of key variables such as GDP growth, unemployment rates, inflation, trade volumes and tax revenue on laundering patterns remains largely unexplored. Advanced analytical techniques like ML algorithms can be leveraged with this "big-picture" economic data to enhance predictive capabilities in this critical, underexplored domain.

METHOD

This study utilises ML algorithms to analyse how macroeconomic factors influence on money laundering risk in Malaysia, following the systematic workflow illustrated in Figure 1. The process begins with the collection of a dataset comprising Malaysian macroeconomic indicators associated with money laundering risk. During data pre-processing, missing values are imputed using the mean of their respective features, data standardisation is applied to ensure consistent scaling across variables, and the dependent variable is transformed into a suitable format for classification.

In the model selection phase, ML algorithms are selected based on their predictive capabilities. Subsequently, the dataset is partitioned into an 80% training set and a 20% testing set to facilitate robust evaluation. The models are then trained to identify patterns between economic indicators and money laundering risks. To optimise results, the workflow includes an iterative loop consisting of hyperparameter tuning (adjusting model settings) and cross-validation (verifying consistency across data subsets).

After each training iteration, a decision point assesses whether the training goal is met, based on predefined performance criteria. If the goal is not achieved, the process returns to hyperparameter tuning and retraining until satisfactory performance is obtained. Once the training objectives are met, the optimised models are applied to the testing phase, where their predictive performance is evaluated on unseen data. This step provides an unbiased assessment of how well the models generalise beyond the training dataset. Model effectiveness is then assessed through performance evaluation, using multiple classification metrics to capture different aspects of predictive accuracy and robustness. Finally, feature importance analysis is conducted to identify the most influential economic variables contributing to the model's predictions. The process concludes after all evaluation and interpretation steps are completed, providing both predictive results and substantive insights into the underlying economic drivers. All model training, testing, and result visualisation were performed using Python software (version 3.13) and its associated library packages, with Matplotlib for plotting.

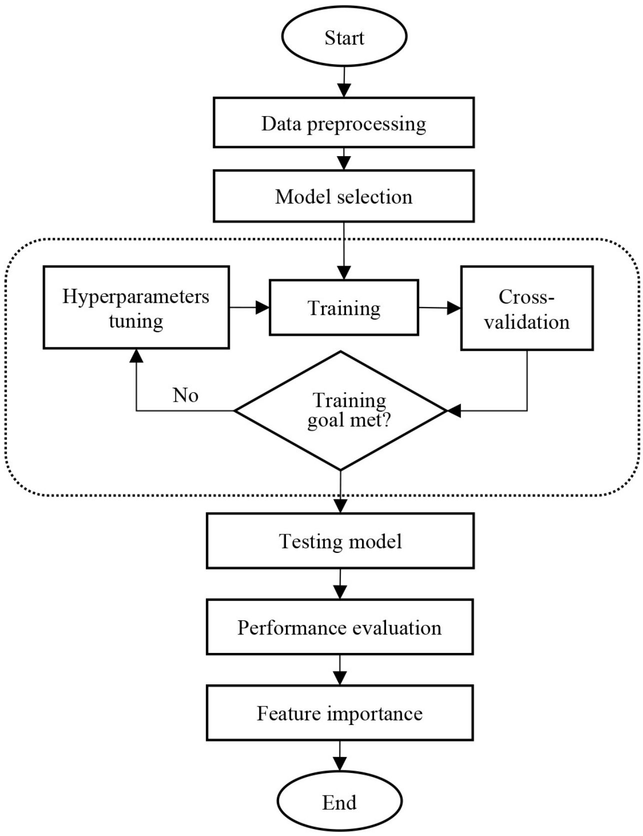


FIGURE 1: Machine learning (ML) process flow chart

Dataset and Variables

Table 1 presents a list of independent variables (i.e., features) representing key economic indicators used to evaluate Malaysia's vulnerability to money laundering risk (dependent variable). These variables were selected for two key reasons. Firstly, they reflect macroeconomic and financial attributes that are widely recognised as determinants of systemic AML risk, many of which contribute indirectly to the Basel AML Index scores. These include:

economic performance (GDP per capita and exchange rate), macroeconomic stability (annual inflation and CPI index), external sector indicators (export and import volume indexes, and the country's net secondary income), and fiscal and social indicators (tax revenue and unemployment rate). Secondly, these indicators sourced primarily from the World Bank and standardised global data repositories provide an objective basis for independently benchmarking a country's AML risk profile in alignment with recent studies (Jullum et al., 2020; Lokanan, 2024a). Taken together, these variables represent multidimensional inputs used to predict a country's money laundering exposure, as measured by its Basel AML risk classification. The dataset spans the period from 2000 to 2024, enabling an insightful view of risk patterns and their association with evolving economic conditions. In this study, the dependent variable is the Basel AML Index Risk, calculated as a country's score on the money laundering. This index, prepared by the Basel Institute on Governance, is an independent, data-driven tool that assesses and ranks countries according to the risk level of money laundering and terrorist financing around the world. The risk score is a composite measure derived from five key domains contributing to money laundering vulnerability: the quality of the AML Framework score, corruption and fraud risks score, financial transparency and standards compliance, public transparency and accountability score, and legal and political risk (Basel Institute on Governance, 2024).

The continuous Basel AML Index scores were transformed into categorical variables using K-Means clustering ($k = 2$). This approach allowed us to address the narrow distribution of scores and the imbalance that fixed, pre-defined thresholds would create. The two clusters represent low-risk (0) and high-risk (1) categories, which were labelled accordingly after ranking their centroids in ascending order. This data-driven approach ensured balanced class representation while preserving the underlying risk patterns in the data. It also aligns with methodologies used in comparable AML research where unsupervised learning is used to discretize risk variables for classification purposes (Lokanan, 2024a). The final dataset was used to train and evaluate ML classification models predicting Malaysia's money laundering risk level based on economic indicators.

TABLE 1
List of variables

Variables	Description	Data source
GDP per capita (current USD)	Gross Domestic Product divided by midyear population; reflects economic development level	Work Bank – World Development Indicators
Exchange rate (LCU per USD, period average)	National currency units per USD; reflects currency stability and international financial exposure	World Bank – World Development Indicators
Inflation, consumer prices (annual %)	Yearly percentage change in consumer prices; measures price stability and economic volatility	World Bank – World Development Indicators
Consumer Price Index (CPI, 2010 = 100)	Relative measure of inflation comparing current prices to a 2010 base year	World Bank – World Development Indicators
Export volume index (2015 = 100)	Measures volume of goods exported, indexed to 2015; indicates trade openness and potential laundering channels	World Bank – International Trade Statistics
Import volume index (2015 - 100)	Measures volume of goods imported, indexed to 2015; indicates exposure to cross-border risks	World Bank – International Trade Statistics
Tax revenue (% of GDP)	Government tax income as a percentage of GDP; reflects fiscal transparency and governance capacity	World Bank – International Trade Statistics
Unemployment, total (% of total labour force) (national estimate)	The percentage of the total labour force that is unemployed but actively seeking employment	World Bank – World Development Indicators

(*Continued on next page*)

TABLE 1 (*Continued*)

Variables	Description	Data source
Net secondary income (Net current transfers from abroad) (current USD)	The net value of current transfers from abroad, including remittances sent by migrants to their home country	World Bank – World Development Indicators
Basel AML Index Risk	Composite risk score (0–10) reflecting a country's exposure to money laundering and terrorist financing, and discretised into low-risk (0) and high-risk (1) categories for classification analysis	Basel Institute on Governance – AML Index Reports

Synthetic Minority Over-sampling Technique (SMOTE) for Imbalanced Data

This study addresses a moderately imbalanced classification problem in which the target variable, Basel AML Index Risk, is unevenly distributed between low-risk (0) and high-risk (1) observations. In the dataset used, high-risk observations constitute approximately 42% of the sample, while low-risk observations account for about 58%. This creates a class imbalance that can bias ML models toward the majority class and reduce their ability to correctly identify minority-class observations (He & Garcia, 2009; Lokanan, 2024a; Phyu & Uttama, 2024). To mitigate this issue, the SMOTE was employed during model training. SMOTE is an over-sampling approach that generates synthetic minority-class instances by interpolating between existing minority observations and their nearest neighbours, thereby improving class balance without simply duplicating data points (Chawla et al., 2002). This approach improves the learning of decision boundaries and supports more reliable predictions by ML algorithms.

Machine Learning (ML) Algorithms Considered

In this study, three well-established classification ML algorithms – RF, LR and NB were applied to predict the money laundering risk in Malaysia based on selected macroeconomic indicators. These models have been successful in earlier research to detect fraud, financial irregularities and illegal activities (Sintayehu & Seid, 2023). They have the flexibility of applying defined risk characteristics such as financial and trade metrics when training and testing models on historical data. In addition, they are also able to generalise patterns of multi-dimensional input data, therefore suiting them in the prediction of country-level money laundering risks and classifying levels based on profile learning (Lokanan, 2024a).

Random Forest (RF)

RF is an ensemble learning approach which generates a set of decision trees and aggregates their output to improve predictive performance. It functions as an ensemble classifier which utilises the method of randomisation for developing a group of independent and diverse decision trees. RF can be defined as:

$$\{k(x,\theta_i), i = 1, 2, \ldots, K\}, \tag{1}$$

Where θ_i is a mutually independent random vector parameter and x is the input data. Each decision tree within the forest is built based on the random vector as a parameter, which determines the random selection of features and a subset of samples used in the training set. By averaging the predictions from these multiple decision trees, it reduces the risk of overfitting, which is generally common with a single decision tree. RF works well for both classification and regression problems.

Naïve Bayes (NB)

NB is a simple classification learning algorithm based on Bayes' theorem. Its fundamental principle relies on a "naïve" but powerful assumption that all input features are independent from one other conditional on the class we attempt to predict. In particular, for an input vector represented by the attribute (or feature) values ($a_{1,\ldots,}a_n$ for attributes $A_i, \ldots, A_n$), the classifier computes the probability of a specific class ($C = c$) according to:

$$P(A_1 = a_{1,\ldots,}A_n = a_n | C = c) = \prod\nolimits_{i=1}^{n} P(A_i = a_i C = c) \quad (2)$$

This "naïve Bayes assumption" greatly simplifies the problem of estimating the joint probability $P(a_1, \ldots, a_n|c)$, turning a potentially exponential and data-intensive calculation into a practical product of individual conditional distributions. This characteristic makes NB well-suited for classification purposes.

Logistic Regression (LR)

LR is a statistical approach used to distinguish between suspicious and legitimate financial transactions in money laundering cases. In these cases, the dependent variable can be binary-coded as '1' (taking it to indicate that a transaction is suspicious) and '0' for a non-suspicious one (Yasaka, 2017). LRs are widely used because of their simplicity, interpretability and capability in modeling for structured financial and economic data (Ghulam & Szalay, 2024; Jullum et al., 2020). Unlike traditional linear regression that predicts an outcome directly, LR estimates the probability that the dependent variable will fall into a particular category, given the context and independent variables.

Let Y be the binary response variable, defined such that $Y_i = 1$ represents the presence of characteristics, $Y_i = 0$ represents uts absence, and the data $[Y_{1,}, Y_2, \ldots, Y_n]$ are independent. Additionally, let $x = (x_{1,}, x_2, \ldots, x_p]$ be the set of explanatory variables which can be discrete, continuous, or mixed types. Then, the logistic function for probability of success, π_i is given by:

$$\text{logit } (\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_o + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i \quad (3)$$

The coefficient β_o is the expected value of Y when all independent variables are zero, and β_{ip} quantifying the change in the log-odds of the dependent variable for a one-unit increase in the independent variable. The independent variable, X_p affects the dependent variable and ε_p is the random error term, which denotes unexplained variance. This equation is the basic structural model for modelling associations between the independent variables and probability of a transaction being classified as suspicious.

Hyperparameter Tuning

Before applying the ML algorithms, hyperparameter tuning was conducted to improve model generalisation and ensure that each algorithm operated under an optimal configuration prior to final evaluation. Table 2 summarises how each ML algorithm was optimised. For the RF and LR models, hyperparameter tuning was performed using GridSearchCV with five-fold cross-validation. Model hyperparameters were optimised by maximising the Receiver Operating Characteristic–Area Under the Curve (ROC–AUC), a threshold-independent metric that evaluates discriminatory performance across all classification thresholds and is robust to class imbalance in AML risk modeling.

For the RF model, three key hyperparameters were systematically tuned: the number of trees in the ensemble (n_estimators), the maximum depth of each decision tree (max_depth), and the minimum number of samples required to split a node (min_samples_split). These parameters directly influence model complexity and generalisation performance. For LR, tuning focused on the regularisation parameter (C), which controls the strength of penalisation applied to the model coefficients. This enables effective management of the bias–variance trade-off and helps prevent overfitting. The solver was fixed to "lbfgs" to ensure numerical stability and efficient convergence. In contrast, NB did not undergo hyperparameter tuning, as it does not possess tunable parameters that significantly affect model complexity or learning behaviour. Its parameters are estimated directly from the data based on distributional assumptions. Consequently, NB was included as a baseline classifier for comparison with the tuned models.

TABLE 2
Hyperparameters tuning by algorithms

Algorithm	Tuning metric	CV strategy	Hyperparameters tuned
Random Forest (RF)	ROC-AUC	5-fold CV (GridSearchCV)	n_estimators ∈ {100, 200}; max_depth ∈ {None, 10, 20}; min_samples_split ∈ {2, 5}
Logistic Regression (LR)	ROC-AUC	5-fold CV (GridSearchCV)	C ∈ {0.1, 1, 10}; solver = lbfgs
Naïve Bayes (NB)	N/A	N/A	N/A

Note: CV = cross-validation; N/A = not applicable

Classification Metrics and Model Performance

To assess the performance of the ML classification models used in this study, a confusion matrix matrix—a standard evaluation tool in ML applications (Lokanan, 2024b)—was employed. In binary classification tasks, the matrix summarises model predictions against actual outcomes across four distinct categories (see Figure 2). True Positive (TP), located in the top-left cell, occurs when the model correctly predicts that a country falls into the high-risk category, and this aligns with the actual classification. Conversely, True Negative (TN), found in the bottom-right cell, represents cases where the model accurately identifies a country as low-risk, consistent with the ground truth. False Positives (FP), in the bottom-left cell, occur when the model incorrectly classifies a low- or medium-risk country as high-risk, potentially leading to unnecessary scrutiny. In contrast, False Negative (FN), located in the top-right cell, represent instances where the model fails to detect a high-risk country, classifying it instead as low-risk.

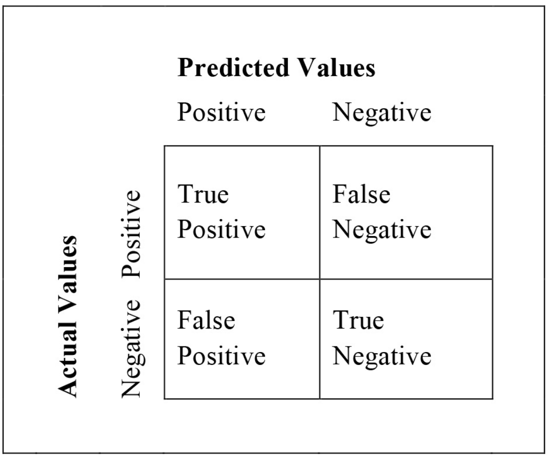


FIGURE 2: Confusion matrix

The elements of the confusion matrix allow for the calculation of several performance metrics that capture different aspects of model behaviour, particularly in the presence of class imbalance. Table 3 summarises the evaluation metrics used in this study, along with their mathematical definitions and descriptions.

TABLE 3
Evaluation metrics

Metrics	Formulae	Description
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	Proportion of correctly predictions (both TP and TN) among the total number of cases.
Balanced accuracy	$\frac{1}{2}\left(\frac{TP}{TP + FN} + \frac{TN}{TN + FP}\right)$	Average of sensitivity and specificity; adjusts for class imbalance.
Precision	$\frac{TP}{TP + FP}$	Proportion of positive predictions that are correct.
Recall (Sensitivity)	$\frac{TP}{TP + FN}$	Proportion of actual positives correctly identified.
F1-Score	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$	Harmonic mean of precision and recall; balances false positives and false negatives.
Cohen's Kappa	$\frac{P_o - P_e}{1 - P_e}$	Measures agreement between predictions and true labels beyond chance; robust to imbalance.
RMSE	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{p}_i - y_i\right)^2}$	Measures average magnitude of prediction or probability estimation error.
PR-AUC	$\int_0^1 \text{Precision}\,(\text{recall})\; d\,(\text{Recall})$	Area under the Precision–Recall curve; emphasises minority-class performance.
ROC-AUC	$\int_0^1 \text{TPR}(FPR)d\,(FPR)$	Measures overall discriminatory ability across all thresholds; threshold-independent.

Notes: RMSE = Root Mean Square Error; PR-AUC = Precision–Recall Area Under the Curve; ROC-AUC = Receiver Operating Characteristic–Area Under the Curve; p_0 = observed agreement; p_e = expected agreement by chance; y_i = true label of observation i; p_i = the predicted probability for the positive class; N = the total number of observations; TPR = True Positive Rate; FPR = False Positive Rate.

RESULT AND DISCUSSION

Empirical Result

Table 4 offers a comprehensive overview of descriptive statistics for various numerical features, providing insights into economic and money laundering risk indicators in Malaysia from 2000 to 2024. The Basel AML Index Risk (0–10 composite score), covering the period from 2012 to 2024, exhibits a mean of 5.40 and a low standard deviation of 0.24. This suggests a relatively consistent, moderate-to-high perceived anti-money laundering risk, ranging from 5.16 to 6.1. For the ML analysis, missing data points for years outside this range were imputed. Economic indicators such as GDP per capita show a notable spread, reflecting variations in economic prosperity, while the exchange rate demonstrates stability around the reported Ringgit-to-major-currency rates. Macroeconomic stability is further illustrated by annual inflation, which includes a minimum of –1.14, possibly reflecting periods of deflation, and a maximum of 5.44, indicating inflationary pressures. The CPI also shows moderate variability. Trade-related metrics like the Export volume index and import volume index reveal similar average levels and dispersions, suggesting balanced trade dynamics. Fiscal and social indicators, tax revenue and unemployment rate, all display relatively low variability. Finally, net secondary income, with a mean of –4.13E+09 and a wide range from –6.88E+09 to –6.46E+08, indicates consistent net outflows, a common characteristic for economies with significant outward remittances or transfers, which can be observed in Malaysia's balance of payments data. This detailed statistical breakdown provides a foundational understanding of the dataset's characteristics, reflecting key economic and financial attributes (Department of Statistics Malaysia, 2025).

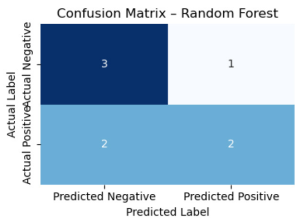
TABLE 4
Descriptive statistics of the numerical features

Variables	Mean	Median	SD	Min	Max
Basel AML Index Risk (0-10)	5.40	5.33	0.24	5.16	6.1
GDP per capita	8567.83	9648.68	2724.31	3943.80	11892.00
Exchange rate	3.77	3.80	0.43	3.06	4.57
Annual inflation	2.09	2.03	1.30	–1.14	5.44
Consumer price index	105.22	104.89	16.55	80.52	132.10
Export volume index	94.49	91.92	18.56	63.66	137.11
Tax revenue	14.03	14.30	1.75	10.88	17.79
Unemployment rate	3.48	3.41	0.43	2.88	4.64
Net secondary income	–4.13E+09	–4.49E+09	1.65E+09	–6.88E+09	–6.46E+08

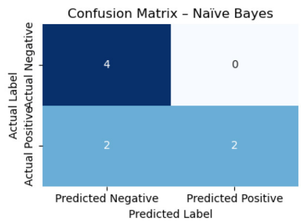
Model Performance

Figure 3 presents the confusion matrices for the RF, NB and LR classifiers used to predict the AML risk classification of Malaysia based on selected economic indicators. For LR, the model correctly classified 2 high-risk observations (TP) and 1 low-risk observation (TN) for Malaysia. However, it incorrectly classified 3 low-risk observations as high-risk (FP) and failed to identify 2 high-risk observations (FN). This indicates that the LR model tends to overestimate AML risk by generating a relatively high number of false positives, while still missing a portion of true high-risk signals derived from the economic indicators. The NB classifier shows strong performance in identifying low-risk conditions for Malaysia, correctly classifying 4 low-risk observations (TN) with no false positives (FP = 0). At the same time, it correctly identified 2 high-risk observations (TP) but misclassified 2 high-risk observations as low-risk (FN). This pattern suggests that NB is conservative in assigning high-risk classifications, effectively avoiding false alarms, but may under-detect some high-risk AML conditions. The RF demonstrates a more balanced classification performance. It correctly classified 2 high-risk observations (TP) and 3 low-risk observations (TN), while producing

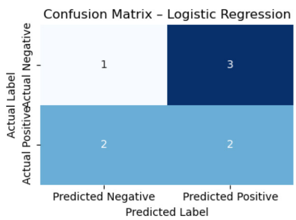
1 FP and 2 FNs. Compared to LR and NB, RF achieves a better trade-off between identifying high-risk and low-risk AML conditions for Malaysia, reducing unnecessary high-risk classifications while maintaining reasonable detection of high-risk cases.



(a) Random Forest (RF)



(b) Naïve Bayes (NB)



(c) Logistic Regression (LR)

FIGURE 3: Confusion matrix for RF, NB and LR

The performance metrics in Table 5 reveal clear differences in the predictive capabilities of the three classification algorithms. Overall, NB exhibits the strongest performance, achieving the highest accuracy and balanced accuracy (0.75), perfect precision (1.0), the highest F1-score (0.6667), and the strongest agreement beyond chance as indicated by Cohen's Kappa (0.50). NB also records the lowest RMSE (0.50) and the highest PR-AUC (0.8110) and ROC-AUC (0.75), highlighting its superior probability estimation and discriminatory power in predicting Malaysia's AML risk. RF demonstrates moderate performance, with accuracy and balanced accuracy of 0.625, a precision of 0.6667, and a Cohen's Kappa of 0.25, indicating reasonable but limited agreement beyond chance; its ROC-AUC (0.5625) suggests modest discriminatory ability. In contrast, LR performs the weakest across all metrics, with low accuracy and balanced accuracy (0.375), a negative Cohen's Kappa (–0.25) indicating agreement worse than chance, the highest RMSE (0.7906), and the lowest PR-AUC (0.3464) and ROC-AUC (0.25). These findings are further supported by Figure 4, which shows that NB consistently lies above the random classification line, RF remains only slightly above it, and LR closely follows or falls below the diagonal, indicating discrimination worse than random classification.

TABLE 5
Performance metric of the classification algorithms

Metric	Random Forest (RF)	Naïve Bayes (NB)	Logistic Regression (LR)
Accuracy	0.625	0.750	0.375
Balanced accuracy	0.625	0.750	0.375
Precision	0.6667	1.0	0.4
Recall	0.5	0.5	0.5
F1-score	0.5714	0.6667	0.4444
Cohen's Kappa	0.25	0.50	–0.25
RMSE	0.6124	0.5	0.7906
PR-AUC	0.7286	0.8110	0.3464
ROC-AUC	0.5625	0.75	0.25

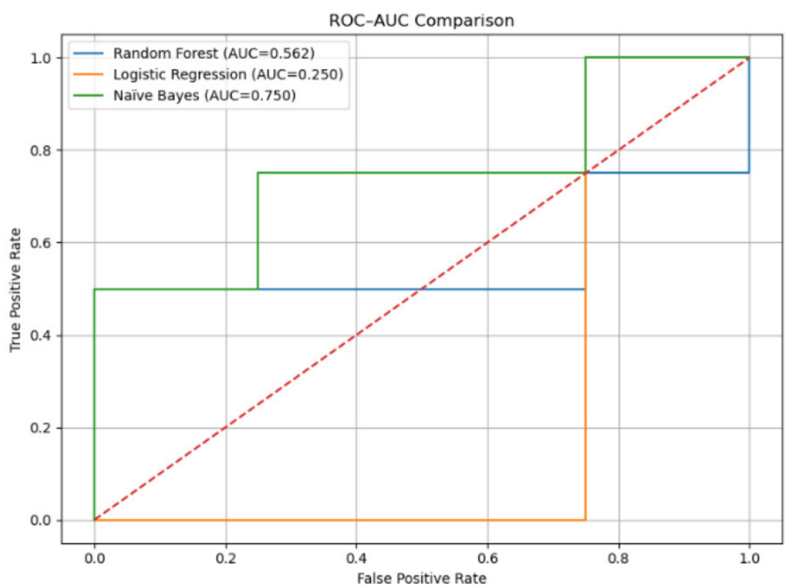


FIGURE 4: ROC-AUC comparison of algorithms

A key explanation for the relatively strong performance of the NB classifier lies in its conditional independence assumption. Although this assumption is often viewed as restrictive, it can be advantageous in high-dimensional macroeconomic settings with limited observations, where complex dependency structures are difficult to estimate reliably. By simplifying the joint probability structure into marginal components, NB reduces estimation variance and helps mitigate overfitting, which is a common concern in small-sample macroeconomic classification tasks (Domingos & Pazzani, 1997; Fernández-Delgado et al., 2014).

The confusion matrix presented earlier further highlights the NB model's cautious prediction behaviour (see Figure 3). The model tends to avoid assigning high-risk labels unnecessarily, while still capturing a portion of high-risk conditions. This reflects a trade-off in which precision is favoured over sensitivity, a characteristic often observed in probabilistic classifiers operating under uncertainty and limited data conditions (Domingos & Pazzani, 1997; Fernández-Delgado et al., 2014). In the context of AML risk assessment, such behaviour may be advantageous from a policy and regulatory perspective, as it helps reduce false alarms that could otherwise result in unwarranted reputational, compliance or economic consequences (Bishop, 2006; Jarmulska, 2022).

In contrast, LR appears more limited in its ability to capture complex relationships between macroeconomic indicators and AML risk, likely due to its reliance on linear decision boundaries. This can result in less stable risk classification in small-sample settings. Meanwhile RF offers a more balanced classification profile; however, its overall discriminatory gains remain modest in low-frequency data environments. This observation is consistent with prior evidence suggesting that ensemble methods may yield limited improvements over simpler classifiers when data are sparse or temporally coarse (Jarmulska, 2022).

Feature Importance

This study employed Recursive Feature Elimination (RFE) to rank the importance of economic variables in predicting money laundering risk in Malaysia. As shown in Figure 5, RFE was applied using the LR algorithm to identify the most influential predictors. The method iteratively removes the least informative features based on the magnitude of the model coefficients, producing a ranked list of variables according to their contribution to classification performance, where a lower rank indicates greater importance.

The results show that the unemployment rate, import and export volume indices, and the exchange rate are the most influential predictors. This suggests that labour market conditions and trade-related factors play a central role in explaining variations in structural money laundering risk. Such findings are consistent with the AML literature, which links economic instability, informal employment, and trade openness to increased exposure to illicit financial flows (Schneider, 2017; Ghulam & Szalay, 2024). The importance of the exchange rate further underscores the relevance of external sector dynamics, as exchange rate movements can affect capital flows, trade misinvoicing and cross-border transactions—well-established channels for money laundering activity (Nguyen, 2013; Walker & Unger, 2009).

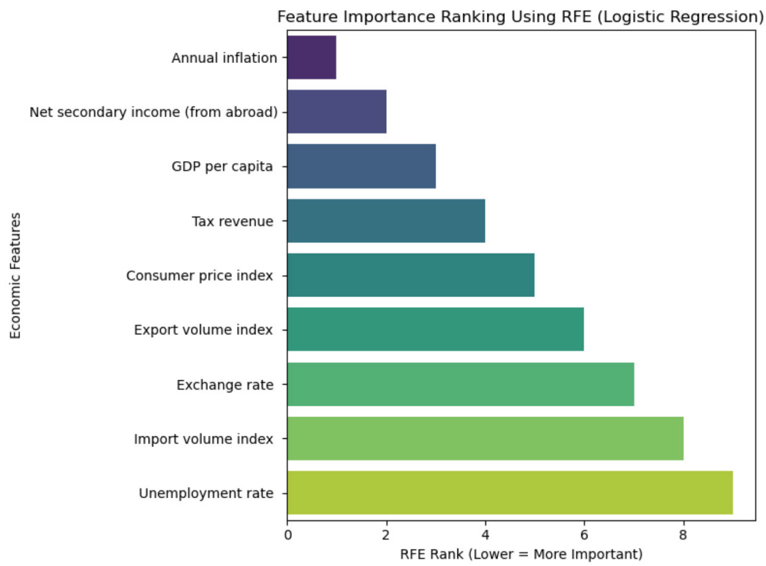


FIGURE 5: Feature importance ranking of economic features

In contrast, variables such as the consumer price index, tax revenue, GDP per capita, net secondary income from abroad, and annual inflation appear less influential in this modelling framework. From a methodological perspective, this future importance analysis emphasises predictors that exhibit stable and consistent associations with the outcome variable. Accordingly, the resulting rankings should be interpreted as relative importance within the specified model, rather than as evidence of causal relationships. Future research may extend this analysis by applying alternative feature selection methods or nonlinear importance measures to further assess robustness.

CONCLUSIONS

This article analysed the impact of key economic indicators on money laundering risk in Malaysia using a comparative ML approach applied to a dataset spanning from 2000 to 2024. Through the application of ML algorithms including RF, LR and NB, the NB emerged as the most effective classifier, consistently demonstrating superior overall performance across multiple evaluation metrics. The confusion matrix and ROC curve analyses further confirmed NB's stronger ability to distinguish between different AML risk levels. The RF showed moderate predictive strength, while LR exhibited weak performance, indicating limited suitability for capturing the underlying complexity of AML risk dynamics in this context. The feature importance analysis conducted through RFE with the LR algorithm provided additional insights into the economic determinants of money laundering risk in Malaysia. The results highlight the unemployment rate, trade-related indicators (import volume) and exchange rate as the most influential factors, suggesting that labour market conditions and external trade dynamics play a critical role in shaping AML risk. These findings offer practical implications for policymakers, financial institutions and regulatory authorities, emphasising the need to closely monitor macroeconomic vulnerabilities and trade-related financial flows when designing targeted and effective AML strategies.

There are several limitations in this study that should be acknowledged. First, the analysis relies on annual macroeconomic indicators, which limits the sample size and temporal granularity; as a result, the ML models capture broad, macro-level patterns in money laundering risk rather than transactional or firm-level dynamics. Second, missing values and class imbalance were addressed using standard preprocessing and resampling techniques, including mean imputation and SMOTE, which may influence model behaviour. Future research could address these limitations by incorporating higher-frequency or transaction-level data, expanding the analysis to multiple countries, and applying network-based or hybrid modeling approaches to enhance predictive accuracy and deepen understanding of money laundering risk dynamics.

ACKNOWLEDGEMENTS

This research was supported by the Short-term Research Grant, Universiti Sains Malaysia [R501-LR-RND002-0000001136-0000].

REFERENCES

Ahmed, J., Mughal, M., & Martínez-Zarzoso, I. (2021). Sending money home: Transaction cost and remittances to developing countries. *The World Economy, 44*(8), 2433–2459. https://doi.org/10.1111/twec.13110

Alotibi, J., Almutanni, B., Alsubait, T., Alhakami, H., & Baz, A. (2022). Money laundering detection using machine learning and deep learning. *International Journal of Advanced Computer Science and Applications, 13*(10), 732–738. https://doi.org/10.14569/IJACSA.2022.0131087

Avdjiev, S., Bruno, V., Koch, C., & Shin, H. S. (2019). The dollar exchange rate as a global risk factor: Evidence from investment. *IMF Economic Review, 67*(1), 151–173. https://doi.org/10.1057/s41308-019-00074-4

Awang Abu Bakar, N., Yahya, N., Idris, N., Ali, E., Zain, J., Khairuddin, E., Abidin, A., Murtaj, S., & Maidin, S. (2024). The determinant factors for the issuance of Central Bank Digital Currency (CBDC) in Malaysia using machine learning framework. *Journal of Applied Data Sciences, 5*(2), 808–821. https://doi.org/10.47738/jads.v5i2.176

Basel Institute on Governance. (2024). Basel AML index 2024. Retrieved 9 August 2025, from https://index.baselgovernance.org

Bishop, C. M. (2006). *Pattern recognition and machine learning.* Springer.

Boateng, N. V., Amoako, N. E. K., Ajay, N. O., & Adukpo, N. T. K. (2025). Harnessing Artificial Intelligence for combating money laundering and fraud in the US financial industry: A comprehensive analysis. *Finance & Accounting Research Journal, 7*(1), 37–49. https://doi.org/10.51594/farj.v7i1.1814

Canhoto, A. (2020). Leveraging machine learning in the global fight against money laundering and terrorism financing: An affordances perspective. *Journal of Business Research, 131*, 441–452. https://doi.org/10.1016/j.jbusres.2020.10.012

Cardoso, M., Saleiro, P., & Bizarro, P. (2022). LaundroGraph: Self-supervised graph representation learning for anti-money laundering [Paper presentation]. *Proceedings of the Third ACM International Conference on AI in Finance*, 130–138. https://doi.org/10.1145/3533271.3561727

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357. https://doi.org/10.1613/jair.953

Department of Statistics Malaysia. (2025). Balance of payments 2025. Retrieved 9 August 2025, from https://www.dosm.gov.my

Domashova, J., & Mikhailina, N. (2021). Usage of machine learning methods for early detection of money laundering schemes. *Procedia Computer Science, 190*, 184–192. https://doi.org/10.1016/j.procs.2021.06.033

Domingos, P., & Pazzani, M. (1997). On the optimality of the simple Bayesian classifier under zero-one loss. *Machine Learning, 29*, 103–130. https://doi.org/10.1023/A:1007413511361

Fallahi, F., Pourtaghi, H., & Rodríguez, G. (2012). The unemployment rate, unemployment volatility, and crime. *International Journal of Social Economics, 39*(6), 440–448. https://doi.org/10.1108/03068291211224937

Fernández-Delgado, M., Cernadas, E., Barro, S., & Amorim, D. (2014). Do we need hundreds of classifiers to solve real-world classification problems? *Journal of Machine Learning Research, 15*, 3133–3181.

FiroozAbadi, S. R., Razmi, M. J., & Bahmani, M. N. D. (2015). Examining the effective social and economic factors on underground economy. *Mediterranean Journal of Social Sciences, 6*(3), 525–534. https://doi.org/10.5901/mjss.2015.v6n3p525

Ghulam, Y., & Szalay, B. (2024). Investigating the determinants of money laundering risk. *Journal of Money Laundering Control, 27*(1), 139–157. https://doi.org/10.1108/JMLC-01-2023-0001

Gupta, A., Dwivedi, D. N., & Jain, A. (2022). Threshold fine-tuning of money laundering scenarios through multi-dimensional optimization techniques. *Journal of Money Laundering Control, 25*(1), 72–78. https://doi.org/10.1108/JMLC-12-2020-0138

Harris, D. A., Pyndiura, K. L., Sturrock, S. L., & Christensen, R. A. G. (2021). Using real-world transaction data to identify money laundering: Leveraging traditional regression and machine learning techniques. *STEM Fellowship Journal, 7*(1), 21–32. https://doi.org/10.17975/sfj-2021-006

He, H., & Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering, 21*(9), 1263–1284. https://doi.org/10.1109/TKDE.2008.239

Hendriyetty, N., & Grewal, B. S. (2017). Macroeconomics of money laundering: Effects and measurements. *Journal of Financial Crime, 24*(1), 65–81. https://doi.org/10.1108/JFC-01-2016-0004

Huong, H., Nguyen, X., Dang, T. K., & Tran-Truong, P. T. (2024). Money laundering detection using a transaction-based graph learning approach [Paper presentation]. *2024 18th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, 1–8. https://doi.org/10.1109/IMCOM60618.2024.10418307

Jarmulska, B. (2022). Random forest versus logit models: Which offers better early warning of fiscal stress? *Journal of Forecasting, 41*(3), 455–490. https://doi.org/10.1002/for.2806

Jayasekara, S. D. (2023). Trade-based money laundering and informal remittance services: Implications on the sustainability of the balance of payments of a small open economy. *Journal of Money Laundering Control, 26*(4), 877–891. https://doi.org/10.1108/JMLC-03-2022-0042

Jensen, R., & Iosifidis, A. (2022). Fighting money laundering with statistics and machine learning. *IEEE Access, 11*, 8889–8903. https://doi.org/10.1109/ACCESS.2023.3239549

Jin, M., & Wang, Z. (2021). Research on the relationships between FATF assessment effectiveness, compliance and money laundering risk [Paper presentation]. *2021 International Conference on Computer, Blockchain and Financial Development (CBFD)*, 533–536. https://doi.org/10.1109/CBFD52659.2021.00113

Jullum, M., Løland, A., Huseby, R. B., Ånonsen, G., & Lorentzen, J. (2020). Detecting money laundering transactions with machine learning. *Journal of Money Laundering Control, 23*(1), 173–186. https://doi.org/10.1108/JMLC-07-2019-0055

Karim, M., Hermsen, F., Chala, S., De Perthuis, P., & Mandal, A. (2024). Scalable semi-supervised graph learning techniques for anti money laundering. *IEEE Access, 12,* 50012–50029. https://doi.org/10.1109/ACCESS.2024.3383784

Ketenci, U. G., Kurt, T., Onal, S., Erbil, C., Akturkoglu, S., & Ilhan, H. S. (2021). A time-frequency based suspicious activity detection for anti-money laundering. *IEEE Access, 9*, 59957–59967. https://doi.org/10.1109/ACCESS.2021.3072114

Khan, A., & Akhtar, N. (2022). A critical appraisal of tax evasion as predicate offence for money laundering. *Pakistan Journal of Multidisciplinary Innovation, 1*(1), 37–44. https://doi.org/10.59075/pjmi.v1i1.32

Khelil, I., Khlif, H., & Achek, I. (2024). The economic consequences of money laundering: A review of empirical literature. *Journal of Money Laundering Control, 27*(5), 901–916. https://doi.org/10.1108/JMLC-09-2023-0143

Kumar, A., Das, S., Tyagi, V., Shaw, R. N., & Ghosh, A. (2021). Analysis of classifier algorithms to detect anti-money laundering. In J. C. Bansal, M. Paprzycki, M. Bianchini, & S. Das (Eds.), *Computationally intelligent systems and their applications* (Vol. 950, pp. 127–140). Studies in Computational Intelligence. Springer. https://doi.org/10.1007/978-981-16-0407-2_11

Kute, D. V., Pradhan, B., Shukla, N., & Alamri, A. (2021). Deep learning and explainable artificial intelligence techniques applied for detecting money laundering: A critical review. *IEEE Access, 9*, 82300–82317. https://doi.org/10.1109/ACCESS.2021.3086230

Labanca, D., Primerano, L., Markland-Montgomery, M., Polino, M., Carminati, M., & Zanero, S. (2022). Amaretto: An active learning framework for money laundering detection. *IEEE Access, 10*, 41720–41739. https://doi.org/10.1109/ACCESS.2022.3167699

Liu, P., & Chen, T. (2024). A genetic algorithm-based method for detecting intermediate accounts in money laundering through analysis of temporal graphs [Paper presentation]. *9th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 263–269. https://doi.org/10.1109/ICSP62122.2024.10743874

Lokanan, M. E. (2019). Data mining for statistical analysis of money laundering transactions. *Journal of Money Laundering Control, 22*(4), 753–763. https://doi.org/10.1108/JMLC-03-2019-0024

Lokanan, M. E. (2024a). Predicting money laundering sanctions using machine learning algorithms and artificial neural networks. *Applied Economics Letters, 31*(12), 1112–1118. https://doi.org/10.1080/13504851.2023.2176435

Lokanan, M. E. (2024b). Predicting money laundering using machine learning and artificial neural networks algorithms in banks. *Journal of Applied Security Research, 19*(1), 20–44. https://doi.org/10.1080/19361610.2022.2114744

Mahmud, A. L. B. & Ismail, A. H. (2023). The effectiveness of anti-money laundering regulations of Malaysian commercial banks. *International Journal of Academic Research in Accounting Finance and Management Sciences, 13*(1), 104–114. https://doi.org/10.6007/IJARAFMS/v13-i1/15853

Masum, M. H., Banik, A., Hasan, M. T., Zolkaflil, S., Mustapha Naz, S. N. F. S., Razali, F. M., & Tarmizi, M. A. (2024). Anti-money laundering and emerging economy: Evidence from Bangladesh. *Journal of Infrastructure, Policy and Development, 8*(6), 3720. https://doi.org/10.24294/jipd.v8i6.3720

Mohan, A., Karthika, P. V., Sankar, P., Maya Manohar, K., & Peter, A. (2023). Improving anti-money laundering in bitcoin using evolving graph convolutions and deep neural decision forest. *Data Technologies and Applications, 57*(3), 313–329. https://doi.org/10.1108/DTA-06-2021-0167

Nazar, S., Raheman, A., & Anwar ul Haq, M. (2024). The magnitude and consequences of money laundering. *Journal of Money Laundering Control, 27*(5), 808–824. https://doi.org/10.1108/JMLC-09-2022-0139

Nazri, S. N. F. S. M., Zolkaflil, S., & Omar, N. (2019). Mitigating financial leakages through effective money laundering investigation. *Managerial Auditing Journal, 34*(2), 189–207. https://doi.org/10.1108/MAJ-03-2018-1830

Nguyen, L. C. (2013). The growing threat of money laundering to Vietnam: The necessary of intensive countermeasures. *Journal of Money Laundering Control, 16*(4), 321–332. https://doi.org/10.1108/JMLC-05-2013-0014

Phillips, J., & Land, K. C. (2012). The link between unemployment and crime rate fluctuations: An analysis at the county, state, and national levels. *Social Science Research, 41*(3), 681–694. https://doi.org/10.1016/j.ssresearch.2012.01.001

Phyu, T. H., & Uttama, S. (2024). Enhancing money laundering detection addressing imbalanced data and leveraging typological features analysis [Paper presentation]. *21st International Joint Conference on Computer Science and Software Engineering (JCSSE)*, 330–336. https://doi.org/10.1109/JCSSE61278.2024.10613724

Reza, F., Islam, M. S., & Anando, K. S. (2024). Money laundering hinders the growth of the economy of a country. *Saudi Journal of Economics and Finance, 8*(01), 9–13. https://doi.org/10.36348/sjef.2024.v08i01.002

Rider, B. (2019). Editorial. *Journal of Financial Crime, 26*(1), 2–4. https://doi.org/10.1108/JFC-09-2018-0103

Sairin, N., Salisi, M., & Bujang, I. (2020). Determining macroeconomic factor of financial distress in Malaysia. *Malaysian Journal of Business and Economics, 2*, 29–36. https://doi.org/10.51200/mjbe.v0i0.2107

Sarigul, H. (2013). Money laundering and abuse of the financial system. *The International Journal of Business and Management, 2*(5), 51–65.

Schlenther, B. (2013). The taxing business of money laundering: South Africa. *Journal of Money Laundering Control, 16*(2), 126–141. https://doi.org/10.1108/13685201311318485

Schneider, F. (2017). The dark side: Crime has gone global. In *A closer look at globalization: The positive facets and the dark faces of a complex notion* (Trilogue Salzburg 2017) (pp. 72–101). Bertelsmann Stiftung.

Schneider, F., & Windischbauer, U. (2008). Money laundering: Some facts. *European Journal of Law and Economics, 26*(3), 387–404. https://doi.org/10.1007/s10657-008-9070-x

Shah, I. H., & Aish, K. (2022). A nexus between corruption, money laundering (ML) and inflation: Evidence from South Asian countries. *Journal of Money Laundering Control, 25*(4), 730–741. https://doi.org/10.1108/JMLC-09-2021-0096

Shen, S., & Dong, X. (2019). The structural relationship between Chinese money supply and inflation based on VAR Model. *Applied Mathematics, 10*(07), 578–587. https://doi.org/10.4236/am.2019.107041

Sintayehu, K., & Seid, H. (2023). Developing anti money laundering identification using machine learning techniques. *Irish Interdisciplinary Journal of Science & Research, 07*(01), 64–74. https://doi.org/10.46759/IIJSR.2023.7110

Summe, K. A. (2000). The battle against money laundering: An examination of US Law, international cooperative efforts and corporate governance issues. *Journal of Money Laundering Control, 3*(3), 236–244. https://doi.org/10.1108/eb027235

Tanzi, V. (1997). Macroeconomic implications of money laundering. In E. U. Savona (Ed.), *Responding to money laundering: International perspectives* (pp. 91–104). Harwood Academic Publishers.

Tasan, M. P., Balasingam, S., Suppiah, K., & Arumugam, D. (2023). Determinants of money laundering: A study among commercial banks in Malaysia. *E3S Web of Conferences 389*, 09032. https://doi.org/10.1051/e3sconf/202338909032

Tiwari, M., Gepp, A., & Kumar, K. (2020). A review of money laundering literature: The state of research in key areas. *Pacific Accounting Review, 32*(2), 271–303. https://doi.org/10.1108/PAR-06-2019-0065

Tran, Q. P. (2024). Is remittance cost a driver of trade misinvoicing? A case study of Vietnam. *Journal of Economics and Development, 26*(4), 362–382. https://doi.org/10.1108/JED-04-2024-0158

UNODC. (n.d.). United Nations Convention Against Illicit Traffic in Narcotic Drugs and Psychotropic Substances (Vienna Convention). United Nations Office on Drugs and Crime. Retrieved 4 May 2025, from https://www.unodc.org

Vagaská, A., Gombár, M., & Korauš, A. (2022). Mathematical modeling and nonlinear optimization in determining the minimum risk of legalization of income from criminal activities in the context of EU member countries. *Mathematics, 10*(24), 4681. https://doi.org/10.3390/math10244681

Walker, J., & Unger, B. (2009). Measuring global money laundering: The Walker gravity model. *Review of Law & Economics, 5*(2), 821–853. https://doi.org/10.2202/1555-5879.1418

Yang, G., Liu, X., & Li, B. (2023). Anti-money laundering supervision by intelligent algorithm. *Computers & Security, 132*, 103344. https://doi.org/10.1016/j.cose.2023.103344

Yasaka, N. (2017). Data mining in anti-money laundering field. *Journal of Money Laundering Control, 20*(3), 301–310. https://doi.org/10.1108/JMLC-09-2015-0041

Yusoff, Y. H., Mohamad Azhar, A. S., Rafidi, F. I., Yunus, N., Azlan, N. J., & Yusop, R. (2023). Money laundering: Factors leading to money laundering in gold investment company in Malaysia. *International Journal of Academic Research in Business and Social Sciences, 13*(5), 940–947. https://doi.org/10.6007/IJARBSS/v13-i5/16599

Zhang, G., Gao, Z., Dong, J., & Mei, D. (2023). Machine learning approaches for constructing the national anti-money laundering index. *Finance Research Letters, 52*, 103568. https://doi.org/10.1016/j.frl.2022.103568

Zhang, Y., & Trubey, P. (2019). Machine learning and sampling scheme: An empirical study of money laundering detection. *Computational Economics, 54*(3), 1043–1063. https://doi.org/10.1007/s10614-018-9864-z