Fraud/Suspicious Transaction Detection: GILB Hackathon Concept Paper

SuchitSainju
Islington College
Kathmndu, Nepal
suchitsainju01@gmail.com

Prasidhika Tiwari
Islington College
Kathmandu, Nepal
line 5: email address or ORCID

Shubham Adhikari
Islington College
Kathmandu, Nepal
subhamadhikari62@gmail.com

Ejen Prapati
Islington College
Kathmandu, Nepal
ejenprajapati0@gmail.com

Abstract. **Fraud, intentional deception for financial gain, has surged with digitalization, intensified by the COVID-19 pandemic. The 2024 Nasdaq Verafin Report estimates $485.6 billion in global fraud losses and $3.1 trillion in illicit funds. In Nepal, FIU-Nepal's 2022/23 report notes Suspicious Transaction Reports doubling to 5,935, driven by money laundering and tax evasion. New frauds like online gambling and crypto scams add complexity. This paper analyzes key fraud types—money laundering, tax evasion, and online gambling/crypto fraud—highlighting red flags like structuring and shell companies. It compares traditional rule-based detection with an AI/ML hybrid pipeline, using supervised (neural networks) and unsupervised (clustering, network analysis) methods on datasets like PaySim to cut false positives, uncover fraud networks, and adapt to evolving threats for better financial security.**

Key words and phrases: **Fraud Detection, Money Laundering, Tax Evasion, Online Gambling, Cryptocurrency Transactions, Suspicious Transaction Reports (STRs/SARs), Financial Crime, Digital Adoption, AI/ML Fraud Detection, Structuring, Money Mules, Shell Companies, Network Analysis, Artificial Neural Networks (ANNs), PaySim Dataset, FIU-Nepal, Global Financial System**

I. Introduction

The Association of Certified Fraud Examiners (ACFE) defines fraud as "any intentional or deliberate act to deprive a person or organization of their property or money by guile, deception, or other unfair means" [1]. Fraud isn't new, historically people have been using various methods to deceive other people. In past, fraud was often localized and had physical presence of the perpetrator.

According to McKinsey & Company Digital adoption has leapfrogged a decade in days during COVID-19, accelerating shift to digital services [2]. As technology integrates into every aspect of life, we increasingly rely on it for our financial needs. This has opened new attack vectors for fraudsters to exploit their target [3]. As digital services have continuously evolved so has the fraudster becoming more advanced and deceptive techniques. The Nasdaq Verafin 2024 Global Financial Crime Report estimates that fraud scams and bank fraud resulted in projected losses of $485.6 billion globally and an estimated $3.1T in illicit funds flowed through the global financial system [4].

In Nepal, there is increasing trend of suspicious transcaction and fraud with Suspisious Transaction Report/Suspicious

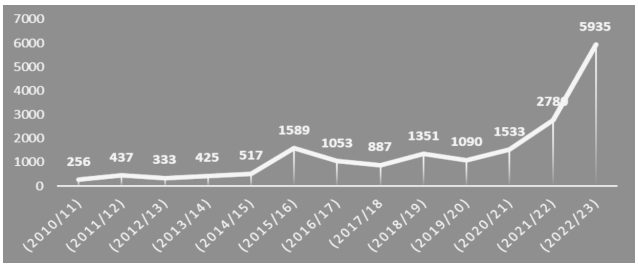

Figure 2: Line chart of STR's/SAR's reported from 2010/11 to 2022/23

Activity Report(STR's/SAR's) received by the FIU Nepal, accorrding to **FIU-Nepal Strategic Analysis Report 2022/23**. According to the report 2022/23 saw a substantial spike in the number of reports received, with 5,935 reports which almost doubled from 2,780 in year 2021/22 [5].

Commercial Banks consistently dominate in reporting, with a substantial spike in 2079/80, totaling 5,187 reports, over two third to the 87% percent, indicative of their significant role in detecting suspicious financial activities.

S.N.	Reporting Institutions (REs)	2073/74 (2016/17)	2074/75 (2017/18)	2075/76 (2018/19)	2076/77 (2019/20)	2077/78 (2020/21)	2078/79 (2021/22)	2079/80 (2022/23)
1	Commercial Banks	949	660	910	924	1403	2380	5187
2	Development Banks	31	23	135	93	64	119	257
3	Finance Companies	0	3	8	1	12	19	44
4	Insurance Companies	4	2	31	4	3	9	19
5	Micro Finance	0	0	0	0	3	44	28
6	Remittance Companies	69	194	263	52	29	187	146
7	Securities Companies	0	2	3	14	18	8	46
8	Cooperatives	0	2	1	2	0	3	1
9	PSD/PSOs	-	-	-	-	-	-	202
10	Government agencies*	0	1	0	0	1	11	5
	Total	**1053**	**887**	**1351**	**1090**	**1533**	**2780**	**5,935**

**Government agencies includes land revenue offices and other government agencies*

Figure 1:STR/SAR's reported by FIU reporting institution

II. FRAUD/SUSPICIOUS TRANSACTION LANDSCAPE ANALYSIS

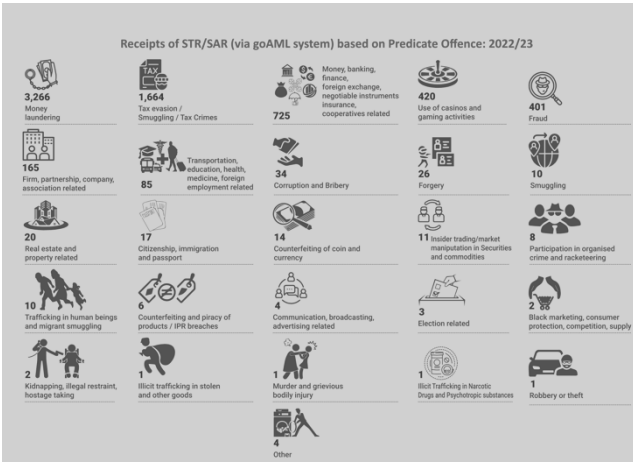

Figure 3: Receipts of STR/SAR 2022/23

A. Money Laundering

According to International Monetary Fund (IMF), money laundering is the processing of assets from criminal activity to obscure their illegal origins *[6]*.

Money laundering involves three steps:

1) Placement: the introduction of the cash into the banking system or legitimate business.

2) Layering: carrying out multiple transactions through multiple accounts with different owners at different financial institutions in the legitimate financial system.

*3) Integration:*merging the funds with money obtained from legitimate activities. *[7]*

According to Nasdaq Verafin 2024 Global Financial Crime Report out of $3.1 trillion illicit funds flowed through the global financial system, money laundering accounted for trillions of dollars. These laundered funds facilitated various criminal activities, including **$782.9 billion** in drug trafficking, **$346.7 billion** in human trafficking, and **$11.5 billion** in terrorist financing [4] . Money laundering was also the most frequently reported STR's/SAR's receipt in Nepal, with a count of 3,266 reports [5].

B. Suspicious Tax Envasion through use of Personal Account

In Nepal, a common form of financial fraud involves the misuse of personal bank accounts for business transactions, primarily to evade taxes and avoid regulatory monitoring. Business owners often conduct high-value transactions through personal savings leading to significant discrepancies between reported and actual earnings.

37% of STR/SARs received (between JNPR2020 to Dec 2022) were related to the suspicious of tax evasion through use of personal bank accounts for business transactions. On an average, monthly 65 such STR/SARs are received by FIU-Nepal during this period. Multiple personal accounts were found to be used for transaction for a single business. For such transactions, personal accounts of proprietor or shareholder is used in most of the cases. Other involved are family members, employees etc.

8% of such business transactions are linked with proprietorship business, 33% with private limited companies and 19% of such firms' registration status is unknown. Nature of such involved businesses are trading, manufacturing, hotels & restaurants, travels, medical etc [5].

C. Online Gambling & Crypto Transaction

Online gambling is banned and illegal in Nepal, yet platforms like **1XBet** and **MostBet** continue to operate using different financial evasion tactics. While many users rely on **VPNs** to hide their IP addresses, a network of agents places bets on their behalf. Regular users transfer money to these agents through **e-wallets, Connect IPS, and e-banking** and other mediums. Then agents funnel the funds through **cryptocurrencies and multiple fake accounts** to obscure the transactions. This process camouflaged betting payouts as personal transactions, allowing them to bypass regulatory monitoring [8].

On **February 3, 2025**, police arrested **24 individuals**, including **10 Indian and 14 Nepali nationals**, for running an illegal online betting operation from **two rented houses in Sanepa, Lalitpur**. The group had reportedly processed **Rs 3.047 billion** in transactions over the past six months [9].

In another case, a syndicate was found to have laundered **Rs 77.787 million** through **eight bank accounts and two fake corporate accounts**, using **cryptocurrency conversions** to evade detection [10].

III. RED FLAGS OF FRAUD DETECTION

A. Structuring/Smurfing

Structuring, also known as smurfing, involves dividing large financial transactions into smaller amounts to evade reporting thresholds set by regulatory authorities. For example, in Nepal, where transactions above NPR 1 million require reporting, an individual might split an NPR 1 million transfer into twenty NPR 50,000 transactions each across different accounts or over several days to avoid detection. This minimizes visibility to financial institutions and regulators, facilitating money laundering and other illicit activities [11].

B. Transaction Patterns

Suspicious transactions often follow recognizable patterns, including:

1) Circular Transfers: Funds move through multiple accounts or entities, returning to their starting point to confuse investigators.

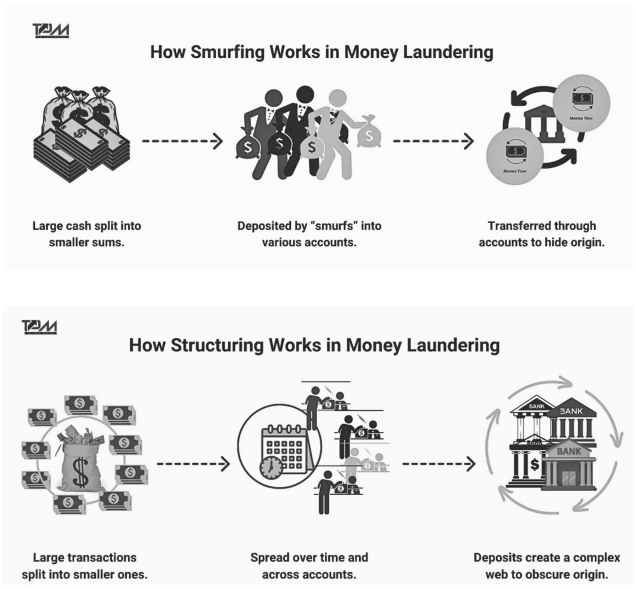

Figure 4: Smurfing/Structuring

2) U-Turn Transactions: Funds exits from a firm and swiftly return via alternate channels, such as foreign remittances or shell company payments.

*3) Rapid Movement***:** Money shifts across several accounts within short timeframes.*[12]*

C. Money Mules

A money mule is someone who, either wittingly or unwittingly, transfers or moves illegally acquired money on behalf of someone else. Mules provide an additional layer of separation between criminals and their proceeds, complicating traceability. Using their own personal bank accounts or opening new ones, they move money through multiple payment channels to confuse regulators of funds and money trail [12].

D. Shell Companies

Shell companies are entities created with no genuine business operations, often used to launder money or conceal ownership. These entities lack physical presence, have vague operational details, and conduct transaction [13].

IV. FRAUD DETECTION APPROCHES

A. Tradtional Fraud Detection Approach

Historically, banks relied on several fundamental techniques to identify fraudulent transactions:

1) Rule-based Systems: Banks established specific thresholds and parameters that would trigger alerts. For example, transactions over certain amounts, multiple transactions in quick succession, or purchases in unusual locations.

*2) Manual Reviews***:** Dedicated fraud analysts would investigate flagged transactions by contacting customers directly to verify legitimacy.

However, these methods face challenges, including generating false positives, which strain resources and negatively impact customer experiences. Siloed data and fragmented data sharing further limit their efficacy against sophisticated fraud schemes.

B. AI/ML enabled Fraud Detection Approach

Banks and Financial Institutions (BFIs) hold vast amounts of user data, including transaction records and personal information. To detect potential fraud, modern approaches leverage **data analytics and predictive modeling**, analyzing these large datasets to uncover suspicious patterns.

Instead of relying only on static rules, **machine learning (ML) models** examine complex relationships between entities, identifying coordinated fraud rings and high-risk accounts. These models detect anomalies in transaction behavior, user activity, or operational irregularities, helping flag potential fraud and generate **detailed risk profiles** for different business areas.

AI and ML are transforming fraud detection by continuously learning from historical fraud patterns and adapting to emerging threats. These systems automatically adjust to new fraud tactics, improving accuracy while reducing false positives. ML algorithms also **prioritize transactions by risk level**, enabling faster responses to the most critical cases.

These models can also be refined incorporating investigator feedback, helping **minimize false alerts**, allowing fraud teams to focus on genuinely suspicious activities.

a.

V. DATASET FOR FRAUD DETECTION

A. Overview of Potential Dataset

For developing and testing our fraud detection pipeline, we require a dataset that captures financial transaction details, user behavior, and indicators of fraudulent activity. Due to the lack of publicly available financial datasets, particularly for money transactions, we reference the PaySim synthetic dataset as an example of the type of data that would be useful. PaySim simulates money transactions with injected fraudulent behavior, providing a realistic foundation for fraud detection. However, since the actual dataset will be provided on the event day, we outline below the types of features and data that would enable us to effectively detect suspicious activities such as money laundering, tax evasion through personal accounts, and online gambling/crypto transactions.

B. Potential Features and Data Dictionary

Below is a data dictionary of potential features, inspired by the PaySim dataset, that we believe are essential for fraud detection. These features are not final but serve as a guide for the type of data the organizers can provide.

Table 1: Data Dictionary of PaySim Dataset

Feature	Description	Data Type
step	Maps a unit of time in the real world; 1 step = 1	Integer

	hour. Total steps = 744 (30 days simulation).	
type	Type of transaction: CASH-IN, CASH-OUT, DEBIT, PAYMENT, or TRANSFER.	Categorical
amount	Amount of the transaction in local currency.	Float
nameOrig	Identifier of the customer who initiated the transaction.	String
oldbalanceOrg	Initial balance of the originating account before the transaction.	Float
newbalanceOrig	New balance of the originating account after the transaction.	Float
nameDest	Identifier of the customer who is the recipient of the transaction.	String
oldbalanceDest	Initial balance of the recipient account before the transaction (unavailable for merchants).	Float
newbalanceDest	New balance of the recipient account after the transaction (unavailable for merchants).	Float
isFraud	Binary label indicating if the transaction is fraudulent (1 = fraud, 0 = non-fraud).	Binary
isFlaggedFraud	Binary label indicating if the transaction is flagged as an illegal attempt (1 = flagged, 0 = not flagged). An illegal attempt is a transfer > 200,000 in a single transaction.	Binary

C. Why this Dataset and Features

The proposed features are critical for our fraud detection pipeline because they enable:

1) Temporal and Behavioral Analysis: Features like step, type, and amount allow us to detect rapid fund movements, structuring, and unusual transaction patterns, which are key indicators of money laundering and online gambling.

2) Network Analysis: nameOrig and nameDest enable the construction of transaction graphs to identify coordinated fraud activities, such as money laundering networks or money mules.

3) Anomaly Detection: Balance-related features (oldbalanceOrg, newbalanceOrig, oldbalanceDest, newbalanceDest) help detect suspicious changes, such as sudden drops or increases, which are common in money laundering and tax evasion through personal accounts.

*4) Supervised Learning:*The isFraud label is essential for training the ANN to classify transactions, while isFlaggedFraud helps prioritize high-risk transactions.

VI. FRAUD DETECTION PIPELINE

A. Proposed Fraud Detection Pipeline

We propose a hybrid fraud detection pipeline combining supervised and unsupervised machine learning to detect suspicious activities like money laundering, tax evasion through personal accounts, and online gambling/crypto transactions in Nepal. The pipeline leverages the potential dataset features and consists of the following stages:

*1) Data Ingestion and Preprocessing***:** Transaction data with features like step, type, amount, nameOrig, nameDest, balance details (oldbalanceOrg, newbalanceOrig, oldbalanceDest, newbalanceDest), isFraud, and isFlaggedFraud.

2) Preprocessing: Clean data, encode type (e.g., CASH-IN, CASH-OUT), and engineer features like transaction velocity (transactions per step) and balance changes (e.g., newbalanceOrig - oldbalanceOrg).

3) Supervised Learning with Artificial Neural Networks (ANNs): Train an ANN using isFraud labels to classify transactions as fraudulent. Use features like amount, type, balance changes, and step to detect known patterns (e.g., structuring, high-value personal account transactions). Output a fraud probability score, flagging transactions (e.g., probability > 0.8) for further investigation.

4) Unsupervised Learning with Network Analysis and Clustering:

i. Network Analysis: Build a graph with nameOrig and nameDest as nodes, using amount and step for edges. Apply Strongly Connected Components (SCC), Smurfing Degree Analysis, PageRank to identify fraud networks (e.g., circular transfers, money mules).

ii. Clustering: Use DBSCAN on features like transaction velocity and amount to group similar behaviors, flagging anomalies (e.g., rapid fund movements).

5) Risk Profiling and Decision Making:
Combine ANN fraud probability, network metrics (e.g., PageRank scores), and clustering anomaly scores into a weighted risk score. Flag high-risk transactions (e.g., risk score > 0.75) for investigation.

6) Alert Generation and Feedback Loop:
Generate alerts for fraud analysts with details like fraud probability and network insights.

Use feedback to retrain the ANN and refine unsupervised models, ensuring adaptability to new fraud tactics.

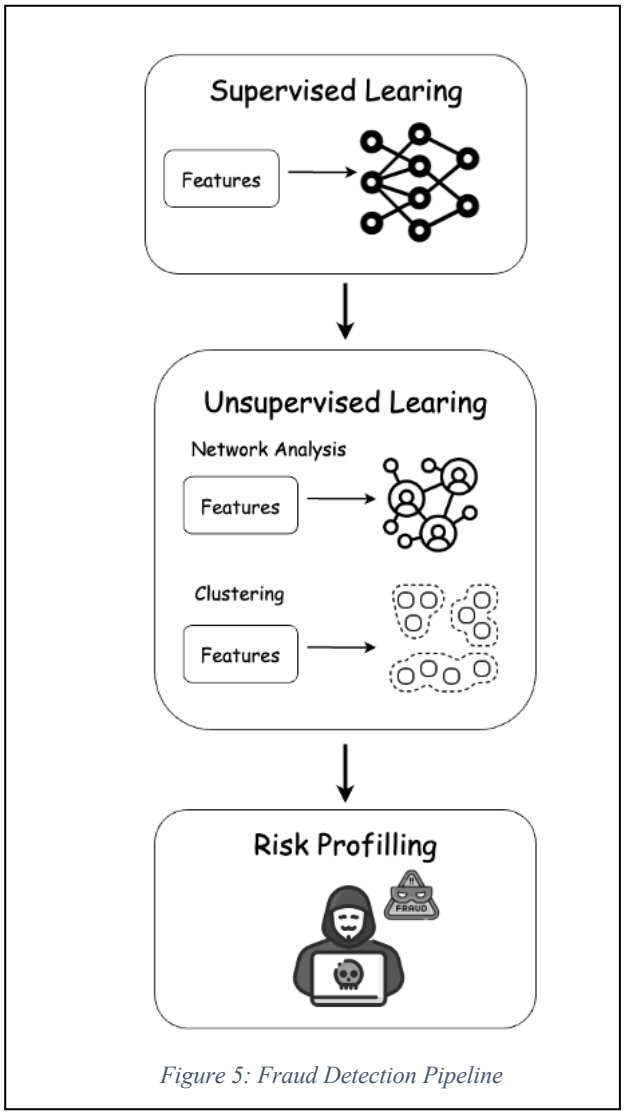

Figure 5: Fraud Detection Pipeline

B. System Architecture

The system architecture for the fraud detection pipeline is designed to be modular, scalable, and efficient, ensuring reliable processing of financial transactions. Below is an overview of the architecture:

1) Data Layer:

The dataset, including transaction data (step, type, amount, etc.), account details (nameOrig, nameDest), and balance information (oldbalanceOrg, newbalanceOrig, etc.).

2) Processing Layer:

i. Feature Engineering: A preprocessing module extracts and transforms features from the dataset (e.g., transaction velocity, balance changes, account age) using Python libraries like Pandas and NumPy.

ii. Supervised Learning (ANN): Implemented using TensorFlow or PyTorch, the ANN model runs on Integrated Graphics Processor (IGP) for training and inference.

iii. Unsupervised Learning: Network analysis and clustering are performed using libraries like NetworkX (for graph analysis with nameOrig and nameDest) and Scikit-learn (for clustering on features like transaction velocity and amount), running on a distributed computing framework like Apache Spark to handle large-scale data.

*3) **Decision Layer**:*

i. Risk Scoring: A custom module integrates the ANN's fraud probability with network analysis metrics (e.g., PageRank scores) and clustering insights (e.g., anomaly scores) to compute a final risk score.

ii. Alert Generation: High-risk transactions are flagged and sent to a fraud investigator including details like the fraud probability, network connections, and clustering results.

*4) **Feedback Layer**:*

Investigator Feedback: A user interface allows fraud analysts to label transactions as fraudulent or legitimate, updating the isFraud labels in the dataset and feeding the results back into the system.

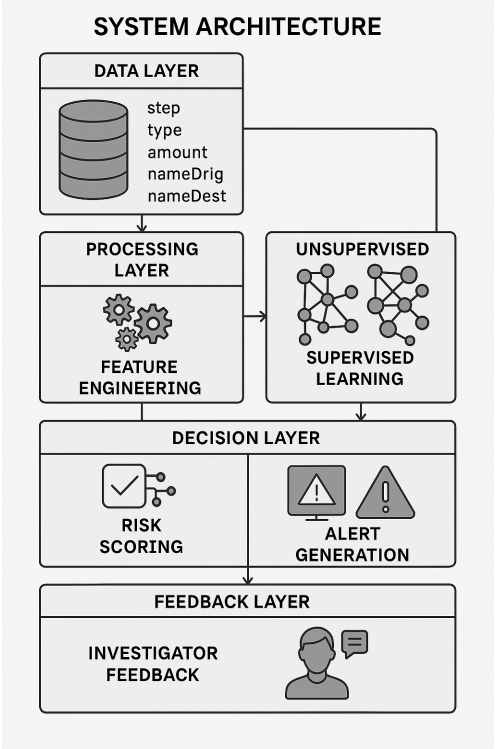

Figure 6: System Architecture for Fraud Detection

C. Expected Outcomes

This pipeline addresses the challenges outlined in the paper by:

1) Reducing false positives

Reducing false positives through the ANN's supervised learning on isFraud labels and the continuous feedback loop.

2) Detecting complex fraud networks

Detecting complex fraud networks (e.g., money laundering, online gambling) using network analysis on nameOrig and nameDest and clustering on features like transaction velocity and amount.

3) Providing actionable insights

Providing actionable insights to fraud investigators through risk scores and detailed reports, including network connections and clustering insights.

4) Adapting to new fraud tactics

Adapting to new fraud tactics by continuously learning from investigator feedback and emerging patterns within transactions.

REFERENCES

[1] "ACFE Fraud Resources," [Online]. Available: https://www.acfe.com/fraud-resources/fraud-101-what-is-fraud.

[2] McKinsey & Company, "A new approach to fighting fraud while enhancing customer experience," *Risk & Resilience,* 2022.

[3] SEON, "SEON-Global Banking Fraud Index 2023," SEON, [Online]. Available: https://seon.io/resources/global-banking-fraud-index/.

[4] Nasdaq, "Nasdaq Verafin 2024 Global Financial Crime Report".

[5] FIU Nepal, "Nepal Rastrya Bank FIU Annual Report 2022/23," NRB, Kathmandu, 2023.

[6] IMF, "IMF Anti-Money Laundering and Combating the Financing of Terrorism (AML/CFT)," [Online]. Available: https://www.imf.org/en/Topics/Financial-Integrity/amlcft.

[7] FATF , "Professional Money Laundering," FATF, 2018.

[8] "FIU Stratigic Analysis Report 2022," NEPAL RASTRA BANK FINANCIAL INFORMATION UNIT, Kathmandu, 2022.

[9] "Rupublica," Republica, 03 02 2025. [Online]. Available: https://www.myrepublica.nagariknetwork.com/news/24-arrested-for-online-betting-operation-with-rs-34-billion-turnover-44-48.html.

[10] "Republica," Republica, 23 03 2023. [Online]. Available: https://myrepublica.nagariknetwork.com/news/six-nepali-agents-make-transaction-of-millions-of-rupees-thru-fake-accounts-in-betting.

[11] "The Perfect Merchant," [Online]. Available: https://theperfectmerchant.com/smurfing-vs-structuring-in-money-laundering/.

[12] "Joint Financial Intelligence Unit HongKong," 2023. [Online]. Available: https://www.jfiu.gov.hk/en/str_screen.html.

[13] FBI, "FBI- Common Fraud and Scams," [Online]. Available: https://www.fbi.gov/how-we-can-help-you/scams-and-safety/common-frauds-and-scams/money-mules.

[14] FATF, "Professional Money Laundering," FATF, 2018.

[15] FIU Nepal, "Nepal Rastrya Bank-FIU Annual Report 2022/23," Nepal Rastrya Bank, Kathmandu, 2023.