LTIMindtree

<Quantum Computing>

<AML and Fraud Detection >

Document Version / Details: Ver. 0.1/ 17-Feb-2023

1. Background

Going into 2023 , we see the regulators agenda under the AML (Anti-Money Laundering) act of 2020 , and the FinCen (Financial Crimes Enforcement Network) have expanded it by its publication of National Proirities , and the first of three final rules on Beneficial Ownership. The FinCen is also going to provide guidance on the advanced use of emerging technologies like quantum computing for imbalanced data sets , Machine learning for transaction monitoring and so on. While there has been focus on various areas in AML , Fraud detection is an area which has been under-regulated.

To create effectual AML processes it is important to move towards a contextual monitoring approach using both internal and external data sources along with innovative technologies to provide the context behind their customers activity towards producing more meaningful alerts and introduce some operational efficiency in the process.

It is critical to move beyond a rules-based approach and apply intelligence-led methods and innovate with latest technologies like Quantum Computing and Machine Learning in Financial Crime and AML.

Here we are trying to map, experiment industry use-case with quantum Machine learning algorithms

2. Goal

Goal : Detection of anomalous transactions in AML and Fraud Detection scenarios for Suspicious Activity Monitoring and False Alert Management where Quantum Machine Learning can be applied.

Use Case : AML and Transaction Monitoring for Fraud detection and anomalous transactions (false negatives) and false positives and SAR (Suspicious Activity Reporting) filings

The Legislation : The Office of Comptroller of Currency and FinCen in the US and various other government entities world-wide have norms related to cash transactions , wire-transfers and credit card transactions for flagging of AML and Suspicious Activity Reporting.

Some of the norms are below :

Office of Comptroller of Currency

- Under the Bank Secrecy Act (BSA), financial institutions are required to assist U.S. government agencies in detecting and preventing money laundering, and:
- Keep records of cash purchases of negotiable instruments;
- File reports of cash transactions and wire transfers exceeding $10,000 (daily aggregate amount); and
- Report suspicious activity that might signal criminal activity (e.g., money laundering, tax evasion).

FinCen SAR - Any transaction conducted or attempted by, at, or through a financial institution that involves or aggregates $5,000 ($2,000 for money services businesses) and the financial institution knows, suspects, or has reason to suspect that the transaction or pattern of transactions of which the transaction is a part:

- Involves funds derived from illegal activity or is intended or conducted to hide or disguise funds or assets derived from illegal activity;
- Is designed, whether through structuring or other means, to evade any relevant requirement of the Bank Secrecy Act;

A FinCEN SAR must be filed within 30 calendar days after the reporting financial institution's initial discovery of information that may provide a basis for filing a report.

3. Fraud Detection : Fraud Categories

The rise of fraud incidents makes it important to learn more about characteristics of datasets associated with different types of frauds. Such understanding will help to implement and better identify the system for fraud detection.

Categorizations and Types of Fraud

1. Structuring and Money Laundering : Make transactions just below the threshold value to avoid detection – e.g. Many $9,900 transactions from the same account
2. Structuring and Money Laundering – e.g Transaction of same customer from different branches on different days with transaction less than $5000 to avoid detection
3. Compromised bank account and zeroing out account amount transactions – One time which empty the bank account
4. Credit card transactions and Money Laundering through large pre-payments and subsequent transactions below threshold
5. Compromised credit card transactions with irrational behavior or purchasing patterns
6. Transactions to a person in the sanctions or criminal or inconsistent with risk categorization and updated profile
7. A user logs in from a different geolocation and device and starts withdrawing large sums of money. Someone is logging in multiple times from a risky IP geolocation pointing to the Cayman Islands.
8. A new user transfers hundreds of small sums to their account and withdraws them in bulk.
9. Money seems to be deposited and withdrawn too quickly.
10. Lastly False alerts – Bank processes transactions flagged above $5000 for operations team to process and to identify them as suspicious transactions while they may be regular transactions slightly above the threshold of $5000

4. Datasets

Dataset Features

A. Transaction and Customer Data

The dataset has 13 input attributes from transaction data and 11 attributes from customer and demographic data. The data also has the historical SAR filings as well with 4 attributes.

The transaction data has the following attributes

1. Instrument or Mode (ACCOUNT , WIRE-TRANSFER , CREDIT-CARD , DEBIT-CARD , CASH , CHEQUE)
2. Type of transaction (PAYMENT , TRANSFER , CASH-IN , CASH-OUT , DEBIT , CREDIT , PRE-PAYMENT , DEPOSIT , LOAN)
3. Amount
4. Account Number of Originator
5. Old Balance
6. New balance
7. Account Number of Destination
8. Old Balance
9. New Balance
10. Above Threshold to be flagged for Operations Team
11. IP Addresses of Originator/Location
12. Is Fraud Marking
13. Time Stamp

Customer or reference data of Originator

1. Bank Identification Number
2. Branch of Account Holder
3. Account Networth
4. AML Risk Rating/Behavioural Profile
5. Name of Account Holder
6. Country or Domicile of Account
7. Onboarding (KYC) Details
8. Beneficiary Details
9. Type of Acount
10. Sanctions List
11. PEP list

Historical SAR filings

1. Name of Account Holder
2. Transaction for which SAR was filed
3. SAR filings Details
4. Date of SAR filing

B. Volume

We are using a dataset of simulated transactions of payments , credit card , debit-card and wire transfer transactions.

The total volume of transactions is 1000 for train-test and 1000 for holdout data.

The data is imbalanced with 40-50 fraudulent transactions for 1000 transactions.

C. Full Dimensionality of DataSet

The dataset has total of 28 attributes and can be split into training and testing data set. The dataset is created with a very few fraud details which make it imbalanced with the number of genuine records relative to the fraudulent ones.

5. Conclusion

Classical machine learning algorithms are currently widespread in use for prediction of fraudulent transactions and fraud detection. QML can support this initiative by increasing the accuracy of prediction for imbalanced data sets.

6. Disclaimer

The material , data and information contained in the document is for general information purposes only from the perspective of the hackathon contest.

While the legislations are mandated and sample of a overall subset of rules , the fraudulent transactions and categories are empirical evidence or based on anecdotal experience.

You should not rely on this data and information for making any business , legal and other decisions.

LTIMindtree

LTIMindtree

Let's get to the future, faster. Together.