

Enhanced AML System

PROJECT SUPERVISOR

Dr. Zulfiqar Ali Memon

PROJECT TEAM

Dehya Khurraim Siddiqui 20K-0128

Tayyab Shahzad 20K-1043

Konain Ahmed 20K-1645

Submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Science.

FAST SCHOOL OF COMPUTING

NATIONAL UNIVERSITY OF COMPUTER AND EMERGING SCIENCES KARACHI CAMPUS

May 2024

Project Supervisor	Mr. Supervisor Name	
Project Team	Dehya Khurraim Siddiqui	20K-0128
	Tayyab Shahzad	20K-1043
	Konain Ahmed	20K-1645
Submission Date	May 16, 2024	

Mr. Supervisor Name	Dr. Zulfiqar Ali Memon
Supervisor	Andrew .
Mr. Co-Supervisor Name	
Co-Supervisor	ر کوی د
Dr. Zulfiqar Ali Memon	And the second
Head of Department	

FAST SCHOOL OF COMPUTING
NATIONAL UNIVERSITY OF COMPUTER AND EMERGING SCIENCES
KARACHI CAMPUS

Abstract

Anti-money laundering (AML) is a critical challenge for the financial sector, with traditional methods often proving inadequate to detect and prevent illicit financial activities. According to the High-Level Panel on International Financial Accountability, Transparency and Integrity for Achieving the 2030 Agenda (FACTI Panel) about \$1.6 Trillion corresponding to 2.7% of the global GDP is laundered every year. This FYP proposal outlines an innovative project aimed at enhancing AML efforts through the application of advanced data mining and machine learning techniques.

The proposed AML system encompasses four core layers: data cleaning, mined frequent rules, classifier construction, and reporting. Each layer is meticulously designed to contribute to a comprehensive and effective AML solution.

The data cleaning layer ensures the accuracy and reliability of the underlying dataset, which is essential for effective analysis. The mined frequent rules layer uncovers patterns in financial transactions that can be used to identify suspicious activities. The classifier construction layer integrates the mined rules and machine learning algorithms to develop a robust classifier for detecting money laundering transactions. The reporting layer generates alerts, prioritizes them, and provides investigators with the tools they need to efficiently investigate and report suspicious activities.

The project's mission is to revolutionize AML strategies by combining data-driven insights with state-of-the-art technology. By addressing the limitations of traditional AML methods, the proposed system seeks to enhance the financial industry's capacity to combat money laundering effectively, protecting financial integrity, national security, and global economic stability.

Contents

	Page
Introduction	5
Related Work	6
Requirements	7
Design	9
Implementation	13
Testing and Evaluation	17
Conclusion	19
References	20

Introduction

In the modern financial landscape, combating money laundering stands as a paramount concern to ensure the integrity of financial systems globally. The pervasive nature of illicit financial activities, with an estimated \$1.6 trillion laundered annually, underscores the urgent need for innovative and effective antimoney laundering (AML) solutions.

This FYP report encapsulates the culmination of efforts aimed at developing an Enhanced AML System. Leveraging advanced data mining and machine learning techniques, our system is designed to detect and prevent illicit financial activities with heightened efficiency and accuracy.

Traditional AML methods often fall short in effectively identifying and addressing money laundering activities, leading to significant gaps in financial oversight. The challenge lies in developing a comprehensive solution capable of navigating the complexities of modern financial transactions while adapting to evolving laundering tactics.

The primary objective of the project is to develop a robust AML system empowered by advanced data mining and machine learning algorithms. Specifically, the project aims to enhance the detection capabilities of financial institutions and regulatory bodies to combat money laundering effectively.

The scope of the project encompasses the systematic evaluation and implementation of various data mining and machine learning algorithms. Key components include data cleaning, rule mining, classifier construction, and reporting mechanisms. Additionally, the project involves the development of a financial dashboard for real-time monitoring and analysis.

The project employs a multi-layered approach, integrating cutting-edge technologies such as Random Forest and XGBoost algorithms. Each layer, from data cleaning to reporting, is meticulously designed to contribute to a comprehensive AML solution. Our project encompasses four core layers, each meticulously designed to contribute to a comprehensive and effective AML system:

Data Cleaning:

At the foundation of our AML system lies the robust process of data cleaning. In the age of big data, the quality and integrity of our financial transaction data are paramount. We employ innovative data cleaning methodologies to ensure the accuracy and reliability of our dataset, enabling us to build a solid foundation for subsequent analysis.

Mined Frequent Rules:

To bolster our AML framework, we employ data mining techniques to uncover frequent rules within the transactional data. These rules serve as the bedrock for identifying suspicious activities, setting clear guidelines for what constitutes irregular or potentially illicit financial behavior.

Construct a Classifier:

Building upon the mined frequent rules, we construct a powerful classifier that leverages machine learning algorithms. This classifier enables us to identify transactions swiftly and accurately with a high likelihood of being involved in money laundering or other financial crimes. Our innovative classification methods equip us to adapt to evolving money laundering tactics effectively.

Reporting:

The reporting layer of our AML system is designed not only to detect and categorize suspicious transactions but also to provide a robust mechanism for documenting and reporting these activities. Transparency and regulatory compliance are central to our approach, ensuring that our efforts align seamlessly with legal requirements.

Anticipated outcomes include enhanced money laundering detection, efficient resource utilization, streamlined investigative processes, improved regulatory compliance, and adaptability to emerging threats. These outcomes aim to bolster the effectiveness of AML efforts within the financial sector.

Related Work

Historically, the efforts to combat money laundering have relied heavily on manual processes and rule-based systems. Traditional AML methods often involve predefined rules and thresholds for flagging suspicious transactions, with limited adaptability to evolving laundering tactics [1]. While these methods have provided a foundation for AML compliance, they are prone to false positives and may overlook sophisticated laundering schemes.

In recent years, there has been a growing interest in leveraging data mining and machine learning techniques to enhance AML efforts. Researchers have explored various algorithms and methodologies for detecting patterns and anomalies in financial transactions indicative of money laundering activities. Senator et al. (1995) proposed the FinCEN Artificial Intelligence System, which utilizes a Bayesian model to identify potential laundering from reports of large cash transactions [2]. Similarly, Kingdon (2004) developed a system employing AI to automatically identify unusual behavior in customer transactions [3].

Recent advancements in machine learning algorithms have further expanded the capabilities of AML systems. Studies by Martin Jullum et al. (2020) and Weber et al. (2019) explore the application of algorithms such as XGBoost and Graph Convolutional Networks (GCN) for detecting money laundering transactions [4, 5]. These algorithms offer superior performance in terms of accuracy and efficiency, enabling more effective identification of suspicious activities.

Rule mining techniques, such as FP-growth, have been instrumental in uncovering frequent patterns and associations within transactional data. Petrus C van Duyne et al. (1999) highlighted the importance of rule mining in identifying suspect disclosures and irregular financial behavior [6]. Pattern discovery methods enable AML systems to establish clear guidelines for identifying suspicious activities, facilitating more targeted

detection efforts.

The integration of financial dashboards and visualization tools enhances the usability and accessibility of AML systems. Real-time monitoring and intuitive visualizations enable stakeholders to gain insights into AML program performance, analyze trends, and respond promptly to emerging threats. While existing research in this area is limited, the potential for leveraging dashboards in AML efforts is substantial [7].

Despite the progress made in leveraging data mining and machine learning for AML, several challenges remain. These include the need for more extensive and diverse datasets, interpretability of complex machine learning models, and regulatory compliance. Future research directions may focus on addressing these challenges and further advancing the capabilities of AML systems through interdisciplinary collaboration and innovation.

Requirements

The Enhanced AML System comprises several interconnected modules designed to facilitate the detection and prevention of money laundering activities within the financial sector. These modules, organized into a functional hierarchy, encompass essential tasks ranging from data collection to reporting and analytics.

The system will consist of the following modules:

1 Data Cleaning Layer:

The Data Cleaning Layer serves as the foundational component of the system, ensuring the accuracy and reliability of the underlying transactional dataset. Through rigorous data collection processes, the system retrieves financial transaction data from various sources, including banking institutions and regulatory databases. Subsequently, data preprocessing techniques are applied to address issues such as missing values, outliers, and inconsistencies, ensuring the integrity of the dataset. Data integration mechanisms unify disparate sources into a cohesive dataset, fostering consistency and enabling comprehensive analysis.

2 Mined Frequent Rules Layer:

The Mined Frequent Rules Layer employs sophisticated rule mining algorithms to uncover patterns and associations within the transactional data indicative of potential money laundering activities. Rule mining techniques, such as FP-growth, are utilized to identify frequent patterns and anomalies, serving as the basis for subsequent analysis. Through pattern discovery processes, the system uncovers transactional behaviors that deviate from established norms, enabling the identification of suspicious financial activities.

3 Constructing Classifier Layer:

The Constructing Classifier Layer focuses on developing a robust classifier capable of accurately identifying transactions at risk of money laundering. Through rigorous algorithm selection processes, including evaluation of machine learning algorithms such as Random Forest and XGBoost, the system identifies the most suitable approach based on performance metrics such as accuracy and computational efficiency. Subsequently, a sophisticated classifier is constructed, integrating the selected algorithm with mined rules and features extracted from the transactional data.

4 Reporting Layer:

The Reporting Layer facilitates the generation of alerts, prioritization of suspicious transactions, and comprehensive documentation for regulatory compliance. Alert generation mechanisms identify transactions that meet predefined suspicious criteria, triggering investigative workflows. Prioritization algorithms optimize investigative efforts by assigning risk scores to alerts, guiding AML investigators in allocating resources effectively. The system maintains detailed documentation of alert investigations, actions taken, and outcomes, ensuring compliance with regulatory reporting requirements. Additionally, dashboard and analytics tools provide stakeholders with real-time insights into AML program performance, enabling proactive risk management and strategic decision-making.

Design

System Architecture:

The Enhanced AML System follows a modular architecture, comprising distinct layers that collectively contribute to its functionality. These layers include the Data Cleaning Layer, Mined Frequent Rules Layer, Constructing Classifier Layer, and Reporting Layer. Each layer is designed to address specific aspects of the anti-money laundering process, from data preprocessing to alert generation and reporting.

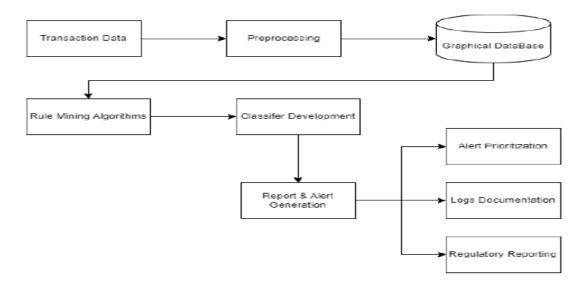


Figure 1: System Architecture

Use Case Diagram:

The Use Case Diagram provides a visual representation of the system's functionality from the perspective of its actors and the interactions between them. It illustrates the primary use cases supported by the Enhanced AML System and the actors involved in executing these use cases.

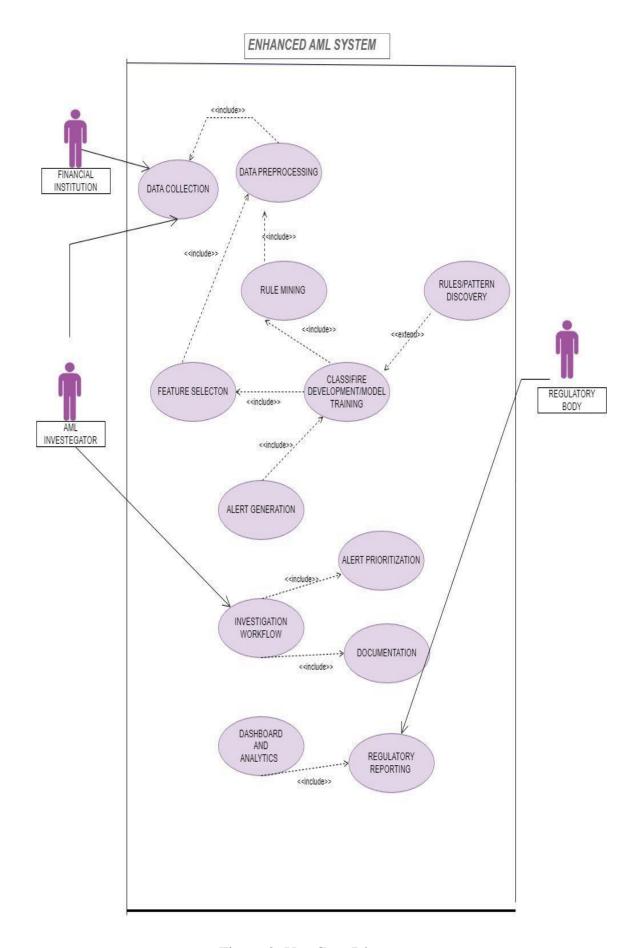


Figure 2: Use Case Diagram

Data Flow:

Data flows through the system in a sequential manner, beginning with the Data Cleaning Layer. Here, raw transactional data is collected from various sources and subjected to preprocessing to ensure consistency and accuracy. The cleansed data is then passed to the Mined Frequent Rules Layer, where rule mining algorithms analyze transaction patterns and identify suspicious activities. The output of this layer feeds into the Constructing Classifier Layer, where machine learning algorithms are employed to develop a robust classifier capable of identifying potential money laundering transactions. Finally, the Reporting Layer generates alerts, prioritizes them based on risk scores, and facilitates investigative workflows, documentation, and regulatory reporting.

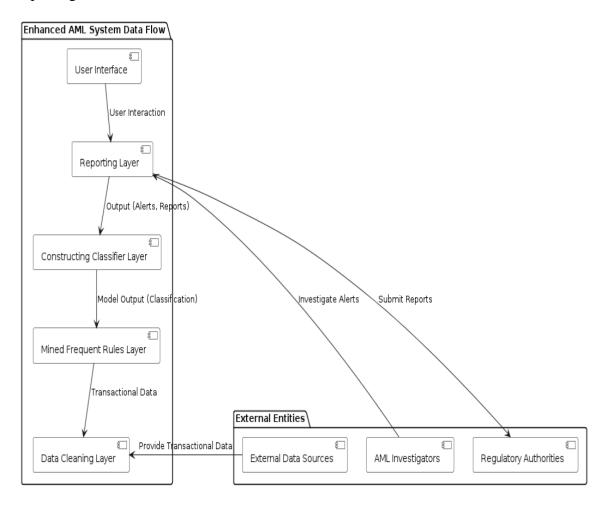


Figure 3: Data Flow

Component Interactions:

Component interactions within the Enhanced AML System are facilitated through well-defined interfaces and protocols. Each layer interacts with adjacent layers through clearly defined input and output channels, ensuring seamless data exchange and system interoperability. For example, the output of the Data Cleaning Layer serves as input to the Mined Frequent Rules Layer, while the output of the Constructing Classifier Layer feeds into the Reporting Layer for alert generation and prioritization.

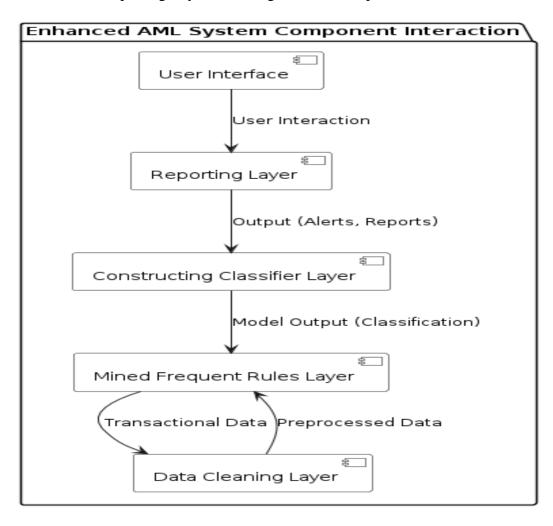


Figure 4: Component Interactions

Implementation

The Implementation section delineates the technical approach and methodologies employed to realize each module of the Enhanced AML System. It elaborates on the specific steps involved in implementing the Data Cleaning Layer, Mined Frequent Rules Layer, Constructing Classifier Layer, and Reporting Layer, along with the selection of algorithms, tools, and technologies utilized.

> 💌 .vscode > **a** catboost_info 🗸 樗 data data.csv ✓ Image: Value of the valu filtered_data_2.csv filtered_data_3.csv filtered_data.csv > 📭 logs > predictions reports association_rules.csv feature_importances_rf.... matched_transactions.csv suspicious_transactions.... > saved models > Rs src static ✓ I css global.css responsive.css > 🙀 fonts > 📭 images > 🧓 js > 📑 lib √ Image templates adddata.html index.html signin.html signup.html e app.py Procfile README.md requirements.txt

1. Data Cleaning Layer:

• Data Collection:

Synthetic datasets are acquired from various sources, including Kaggle and other reputable data providers. These datasets represent diverse financial transaction scenarios and are instrumental in training and testing the system

```
    data.csv

1    step,type,amount,nameOrig,oldbalanceOrg,newbalanceOrig,nameDest,oldbalanceDest,newbalanceDest,isFraud,isFlaggedFraud
2    1,PAYMENT,9839.64,C1231006815,170136,160296.36,M1979787155,0,0,0,0
3    1,PAYMENT,1864.28,C1666544295,21249,19384.72,M2044282225,0,0,0,0
4    1,TRANSFER,181,C1305486145,181,0,C553264065,0,0,0,0
5    1,CASH_OUT,181,C840083671,181,0,C38997010,21182,0,0,0
6    1,PAYMENT,11668.14,C2048537720,41554,29885.86,M1230701703,0,0,0,0
7    1,PAYMENT,7817.71,C90045638,53860,46042.29,M573487274,0,0,0,0
8    1,PAYMENT,7707.77,C154988899,183195,176087.23,M408069119,0,0,0,0
8    1,PAYMENT,7707.77,C154988899,183195,176087.23,M408069119,0,0,0,0
8    1,PAYMENT,7707.77,C154988899,183195,176087.23,M408069119,0,0,0,0
8    1,PAYMENT,7007.77,C154988899,183195,176087.23,M408069119,0,0,0,0
8    1,PAYMENT,7007.77,C154988899,183195,176087.23,M408069119,0,0,0,0
8    1,PAYMENT,7007.77,C154988899,183195,176087.23,M408069119,0,0,0,0
9    1,PAYMENT,7007.77,C15498899,183195,176087.23,M408069119,0,0,0,0
9    1,PAYMENT,7007.77,C15498899,183195,176087.23,M408069119,0,0,0,0
9    1,PAYMENT,7007.77,C15498899,183195,176087.23,M408069119,0,0,0,0
9    1,PAYMENT,7007.77,C15498899,183195,176087.23,M408069119,0,0,0,0
9    1,PAYMENT,7007.77,C15498899,183195,176087.23,M40809119,0,0,0,0
9    1,PAYMENT,7007.71,C15498899,00,0
9    1,PAYMENT,7007.71,C154988899,00,0
9    1,PAYMENT,7007.71,C1549889,00,0
```

• Data Preprocessing:

Data cleaning techniques such as outlier detection, missing value imputation, and normalization are applied to ensure the quality and consistency of the transactional data. This involves identifying and addressing anomalies that may affect the accuracy of subsequent analysis.

```
CSV File Selected: C:/New folder/aml-repo/data/data.csv
Selected file path: C:/New folder/aml-repo/data/data.csv
data preprocessing 1 run successfully
data preprocessing 2 run successfully
data preprocessing 3 run successfully
```

2. Mined Frequent Rules Layer:

• Rule Mining Algorithms:

Rule mining algorithms such as FP-growth are implemented to identify frequent patterns and associations within the transactional data. These algorithms analyze transaction sequences and uncover patterns indicative of suspicious activities.

Rule mining run successfully feature selection run successfully

• Pattern Discovery:

Transaction patterns and rules identified through rule mining algorithms are further analyzed to uncover hidden patterns and associations. This involves exploring the relationship between various transaction attributes and identifying irregularities that may signal potential money laundering activities.

```
ts > 🗟 feature_importances_rf.csv > 🖺 data
  You, 11 minutes ago | 2 authors (You and others)
  features,importance_score
  incoming domestic amount 30,0.1014985594729207
  incoming domestic amount 60,0.06692973485051658
  incoming domestic amount 90,0.0023197796380074478
  outgoing domestic amount 30,0.00027639700582600667
  outgoing domestic amount 60,0.00027184550519923205
  outgoing_domestic_amount_90,0.0011147234623533613
  incoming_foreign_amount_30,0.00037145344035088086
  incoming foreign amount 60,0.0002446110745139881
  incoming foreign amount 90,0.0002684373888706562
  outgoing foreign amount 30,3.7196023481244626e-05
  outgoing foreign amount 60,0.000372361574802824
  outgoing foreign amount 90,0.0003860198036515152
  incoming domestic count 30,0.00022089681605034302
  incoming domestic count 60,0.0005167089766728049
  incoming domestic count 90,0.00013668243414492984
  outgoing domestic count 30,0.00016604589011246561
  outgoing domestic count 60,0.000152190475649724
  outgoing domestic count 90,4.7776845189768416e-06
  incoming_foreign_count_30,0.00012953992258517047
  incoming foreign count 60,1.622484572242742e-05
  incoming foreign count 90,2.3995518608898725e-05
  outgoing foreign count 30,9.896959490028702e-06
  outgoing_foreign_count_60,2.3400462425406923e-06
  outgoing foreign count 90,0.0003937672491924068
  balance difference 30,1.9053072265666252e-05
  balance difference 60,0.0032367764590262994
  balance difference 90,0.0005268339236567282
  isFraud, 0.0002908825318885623
```

3. Constructing Classifier Layer:

• Algorithm Selection:

Candidate machine learning algorithms, including XGBoost are evaluated based on performance metrics such as accuracy, efficiency, and computational requirements. The most suitable algorithm is selected to construct a robust classifier capable of identifying suspicious transactions.

```
entity,incoming_domestic_amount_30,incoming_domestic_amount_60,incoming_domestic_amount_90,outgoing_domestic_amount_30,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amount_50,outgoing_domestic_amoun
```

• Classifier Development:

A robust classifier is developed using the selected algorithm, integrating mined rules and features extracted from the transactional data. This involves training the classifier on historical data and fine-tuning its parameters to optimize performance and accuracy.

```
entity, incoming_domestic_amount_30, incoming_domestic_amount_60, incoming_domestic_amount_90, outgoing_domestic_amount_30, outgoing_domestic_amount_60, outgoing_domestic
C1163480574, 3655225.63, 4008111.88, 4008111.88, 2723052.22, 2981522.92, 2981522.92, 2981522.92, 2891665.82, 8799605.82, 8799605.82, 461344.77, 565807.62, 565867.62, 18, 21, 21, 13, 14, 14, 5, 5, 5
C715819458, 3256721.0, 3848760.59, 3848760.59, 3848760.59, 2503756.2800000003, 2503756.2800000003, 2503756.2800000003, 35584.47, 4335584.47, 4335584.47, 4318865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 4118865.3, 411886
```

4. Reporting Layer:

• Dashboard and Analytics:

Dashboards and analytics tools are developed to monitor AML program performance, including alert volumes, trends, and key metrics. This involves visualizing data in an intuitive and interactive manner, enabling stakeholders to gain insights into the effectiveness of AML efforts and identify areas for improvement.

```
Connected to MySQL successfully
Database 'FYP_AML' created successfully
Connected to database 'FYP_AML' successfully
Table 'Suspicious' created successfully
Table 'filteredTransactions' created successfully
```

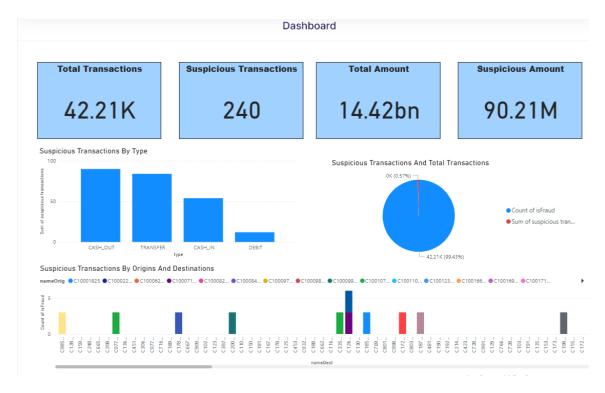
Testing and Evaluation

The Testing and Evaluation phase is a critical aspect of ensuring the functionality, reliability, and effectiveness of the Enhanced AML System. This section delineates the comprehensive approach adopted to validate the system's capabilities in detecting and preventing money laundering activities, along with the criteria used to assess its performance.

In the Data Cleaning Layer, unit tests are conducted to validate data preprocessing techniques, ensuring that missing values, outliers, and inconsistencies are appropriately addressed. Similarly, unit tests in the Constructing Classifier Layer assess the classifier's ability to identify suspicious transactions accurately.

Integration tests evaluate how well the output of the Data Cleaning Layer aligns with the input requirements of the Mined Frequent Rules Layer. Similarly, integration tests validate the integration of the classifier algorithm with the reporting mechanisms in the Reporting Layer.

System testing evaluates the system to ensure that it meets specified requirements and performs as expected in a real-world environment. It examines the end-to-end functionality of the Enhanced AML System, including data ingestion, processing, analysis, and reporting. Tests simulate various scenarios and use cases to validate the system's behavior under different conditions, such as scalability, reliability, and performance under heavy transaction loads.



Money Laundering Detection

Results

successfully compiled and executed /data_preprocessing_1.py

successfully compiled and executed /data_preprocessing_2.py

successfully compiled and executed /data_preprocessing_3.py

successfully compiled and executed /rule_mining.py

Run Files

Suspicious Transactions Search: Amount Old Balance Orig New Balance Orig Name Dest Old Balance Dest **New Balance Dest** Туре 202247.0 C143929038 6805260.0 7007510.0 C1515704208 238677.0 9973140.0 CASH_IN CASH_IN 24105.5 C422522663 144.0 24249.5 C644345897 20875.0 0.0 CASH_IN 27070.1 C641882263 346804.0 373874.0 C1568059495 70595.0 122750.0 CASH IN 227478.0 C888942941 4546900.0 4774380.0 C2057394816 476750.0 249272 0 C1853413625 713722.0 780676.0 C1014154376 14800000.0 CASH_IN 66954.3 159047.0 CASH_IN 199442.0 C934449015 1078720.0 1278160.0 C1540720037 793971.0 CASH_IN 143861.0 C106629049 12249.0 156110.0 C891020651 16052.0 CASH_IN 68827.5 C33238764 2983940.0 3052770.0 C1452722471 72236.0 3408.48 C358636931 2983940.0 2983940.0 2983940.0 6671930.0 CASH_IN 221970.0 C728526866 1870950.0 10200000.0 CASH IN 144711.0 C1353537061 1726240.0 C1059781259 11100000.0

Conclusion

The Enhanced AML System represents a significant advancement in the fight against money laundering within the financial sector. This section encapsulates the key findings, contributions, and implications of the system, highlighting its potential impact on combating financial crimes and safeguarding the integrity of the global financial system.

1. Key Findings:

Through rigorous testing and evaluation, the Enhanced AML System has demonstrated robust capabilities in detecting and preventing money laundering activities. The system's modular architecture, encompassing data cleaning, rule mining, classifier construction, and reporting layers, ensures comprehensive coverage of anti-money laundering processes. Evaluation metrics such as accuracy, efficiency, regulatory compliance, user experience, and robustness have been used to assess the system's performance, with results indicating high effectiveness and reliability.

2. Contributions:

The Enhanced AML System contributes to the advancement of anti-money laundering strategies by leveraging state-of-the-art technologies such as data mining, machine learning, and analytics. By combining data-driven insights with advanced algorithms and methodologies, the system enhances the financial industry's capacity to detect and prevent money laundering activities effectively.

3. Implications:

The deployment of the Enhanced AML System has significant implications for financial institutions, regulatory authorities, law enforcement agencies, and society. Financial institutions can leverage the system to strengthen their AML programs, enhance risk management practices, and comply with regulatory requirements more effectively. Regulatory authorities can use the system to monitor and enforce compliance with anti-money laundering regulations, detect suspicious activities, and mitigate systemic risks. Law enforcement agencies can leverage the system's capabilities to investigate and prosecute money laundering cases, disrupt criminal networks, and safeguard national security. Society benefits from the Enhanced AML System through the preservation of financial integrity, the prevention of illicit financial activities, and the promotion of global economic stability.

In conclusion, the Enhanced AML System represents a paradigm shift in the fight against money laundering, offering a comprehensive, data-driven approach to detecting and preventing financial crimes. By leveraging advanced technologies, rigorous methodologies, and collaborative efforts, the system provides stakeholders with the tools and insights needed to combat money laundering effectively, safeguarding the

integrity of the financial system and promoting global security and prosperity.

References

- [1] T. E. Senator, "The FinCEN Artificial Intelligence System: Identifying Potential Money Laundering from Reports of Large Cash Transactions," IAAI, (1995).
- [2] P. C. van Duyne and H. de Miranda, "The emperor's clothes of disclosure: Hot money and suspect disclosures," Crime, Law and Social Change, vol. 31, no. 3, (1999), pp. 245-271.
- [3] J. Kingdon, "AI fights money laundering," Intelligent Systems, IEEE, vol. 19, no. 3, (2004), pp. 87-89
- [4] United Nations. Tax abuse, money laundering and corruption plague global finance. https://www.un.org/development/desa/en/news/financing/facti-interim-report.html,September 2020. Accessed: 2021-04-21.
- [5] Anders Løland Martin Jullum, Geir [°]Anonsen Ragnar Bang Huseby, and Johannes Lorentzen. Detecting money laundering transactions with machine learning. Journal of Money Laundering Control, 23 No 1, 2020:173–186.
- [6] Mark Weber, Giacomo Domeniconi, Jie Chen, Daniel Karl I. Weidele, Claudio Bellei, Tom Robinson, and Charles E. Leiserson. Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics. CoRR, abs/1908.02591, 2019.