sciendo INTERNATIONAL JOURNAL ON SMART SENSING AND INTELLIGENT SYSTEMS

Article | DOI: 10.2478/ijssis-2024-0024 Issue 1 | Vol. 17 (2024)

Navigating the Complexity of Money Laundering: Anti–money Laundering Advancements with AI/ML Insights

Hitarth Gandhi[1], Kevin Tandon[1], Shilpa Gite[1,2,*], Biswajeet Pradhan[3,*] and Abdullah Alamri[4]

[1]Artificial Intelligence and Machine Learning Department, Symbiosis Institute of Technology, Symbiosis International (Deemed) University, Pune 412115, India

[2]Symbiosis Centre of Applied AI (SCAAI), Symbiosis International (Deemed) University, Pune 412115, India

[3]Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), School of Civil and Environmental Engineering, Faculty of Engineering and IT, University of Technology Sydney, NSW 2007, Australia

[4]Department of Geology and Geophysics, College of Science, King Saud University, Riyadh, Saudi Arabia

*E-mails: biswajeet.pradhan@uts.edu.au, shilpa.gite@sitpune.edu.in

Received for publication
April 04, 2024.

Abstract

This study explores the fusion of artificial intelligence (AI) and machine learning (ML) methods within anti–money laundering (AML) frameworks using data from the US Treasury's Financial Crimes Enforcement Network (FinCEN). ML and deep learning (DL) algorithms—such as random forest classifier, elastic net regressor, least absolute shrinkage and selection operator (LASSO) regression, gradient boosting regressor, linear regression, multilayer perceptron (MLP) classifier, convolutional neural network (CNN), random forest regressor, and K-nearest neighbor (KNN)—were used to forecast variables such as state, year, and transaction types (credit card and debit card). Hyperparameter tuning through grid search and randomized search was used to optimize model performance. The results demonstrated the efficacy of AI/ML algorithms in predicting temporal, spatial, and industry-specific money-laundering patterns. The random forest classifier achieved 99.99% average accuracy in state prediction, while the gradient boosting regressor and random forest classifier excelled in predicting year and state simultaneously, and credit card transactions, respectively. MLP and CNN showed promise in the context of debit card transactions. The gradient boosting regressor performed competitively with low mean squared error (MSE) (2.9) and the highest *R*-squared (R^2) value of 0.24, showcasing its pattern-capturing proficiency. Logistic regression and random forest classifier performed well in predicting credit card transactions, with area under the receiver operating characteristic curve (ROC_AUC) scores of 0.55 and 0.53, respectively. For debit card prediction, MLP achieved a precision of 0.55 and recall of 0.42, while CNN showed a precision of 0.6 and recall of 0.54, highlighting their effectiveness. The study recommends interpretability, hyperparameter optimization, specialized models, ensemble methods, data augmentation, and real-time monitoring for improved adaptability to evolving financial crime patterns. Future improvements could include exploring the integration of blockchain technology in AML.

Keywords

anti-money laundering, AI/ML, FinCEN dataset, USA

I. Introduction

In today's financial landscape, the threat of financial crimes, especially money laundering, looms large, necessitating a paradigm shift in the approach to anti–money laundering (AML) legislation. In an era where financial crimes, especially money laundering, loom as pervasive threats, traditional AML strategies grapple with the relentless evolution of illicit tactics, highlighting an urgent need for innovative solutions. The core motivation behind this research stems from the escalating complexity of financial crime patterns and the limitations of existing AML methodologies in effectively combating these threats. While traditional approaches provide foundational frameworks, they often fall short in addressing the dynamic nature of money-laundering schemes. Consequently, there is a critical gap in understanding the intricate interplay between evolving money-laundering tactics and the regulatory responses required to thwart such illicit activities effectively.

This research addresses several critical research gaps in the realm of AML frameworks and artificial intelligence (AI)/machine learning (ML) integration. One notable gap is the interpretability of AI/ML models, particularly in the context of financial crime detection. By showcasing the effectiveness of various algorithms such as random forest and gradient boosting in identifying money-laundering patterns, this study contributes to the ongoing discourse on improving the interpretability of these models for regulatory compliance and risk assessment in AML frameworks. Through a meticulous analysis of a rich dataset sourced from the Financial Crimes Enforcement Network (FinCEN) of the US Treasury, the study embarks on a journey through the labyrinth of financial crime patterns. As the narrative unfolds, it accentuates the escalating urgency for innovative solutions, highlighting the inherent limitations of conventional AML methodologies. Empowered by empirical insights, this exploration unearths nuanced trends, revealing both vulnerabilities and strengths within the current AML framework. Moreover, it ventures into the realm of practical implications, illuminating the path forward for AML practitioners, policymakers, and regulatory bodies. Yet, amid the promise of AI and ML, this narrative also candidly confronts the challenges and constraints, providing a balanced perspective on their real-world applicability in the relentless battle against financial crime. The study encourages AML professionals and regulatory bodies to adopt advanced technologies and methodologies to keep pace with evolving money-laundering tactics. This cultural shift toward embracing AI/ML technologies in AML frameworks promotes a proactive approach to financial crime prevention, leading to more robust and adaptive systems that can effectively counter sophisticated illicit activities.

Money laundering typically involves a process of disguising the origins of illegally obtained money, making it appear legitimate [1,2]. This process generally consists of three main stages or layers. Placement is the first stage of money laundering and involves introducing illegally obtained funds into the legitimate financial system. This can be done through various means, such as depositing cash into bank accounts, purchasing assets such as real estate or luxury goods, or using money transfer services. The objective is to break the link between the illicit funds and their criminal origins. Criminals often divide large sums of money into smaller amounts to avoid detection, a process known as "smurfing." They may also use techniques such as "structuring," which involves making multiple deposits or transactions in amounts below reporting thresholds to avoid suspicion.

Layering is the second stage of money laundering and involves distancing the illicit funds from their original source through a series of complex financial transactions. The goal is to obscure the paper trail and make it difficult for law enforcement agencies to trace the funds back to their criminal origins. Techniques used in layering include transferring funds among multiple accounts, making international wire transfers, purchasing and selling assets across different jurisdictions, and engaging in transactions with shell companies or offshore entities.

Integration is the final stage of money laundering and involves reintroducing the laundered funds into the legitimate economy in such a way that they appear to be derived from lawful activities. This involves investing the funds in legitimate businesses, purchasing additional assets, or using them to finance legitimate projects. Once the laundered funds have been integrated into the legitimate economy, they can be freely spent or reinvested without raising suspicion. At this stage, the funds appear to have a legitimate source, making it difficult for law enforcement agencies to distinguish them from legally obtained money [3,4].

Integration effectively "cleans" the illicit funds, allowing criminals to enjoy the proceeds of their illegal activities without attracting attention or facing legal consequences. To detect and prevent such illicit financial activities, modern AML systems use a

combination of advanced technical mechanisms and methodologies. Key techniques include transaction monitoring systems, which analyze financial transactions in real time to identify suspicious patterns; risk-based approaches, which prioritize resources based on the risk profiles of customers and transactions; and ML algorithms, which enhance the detection of anomalies by learning from historical data and identifying complex patterns indicative of money laundering. Additionally, network analysis is a powerful tool for identifying and understanding the intricate relationships and interactions between entities involved in financial transactions. By constructing a network wherein nodes represent entities (such as individuals, accounts, or institutions) and edges represent financial transactions, network analysis provides a visual and analytical means to detect and interpret patterns indicative of money laundering. Different techniques are utilized to uncover relationships and connections among entities that comprise money-laundering networks. These methods are complemented by robust customer due diligence (CDD) processes, which ensure thorough background checks and continuous monitoring of customer activities. By integrating these sophisticated tools and methodologies, AML systems strive to stay ahead of evolving financial crimes and safeguard the integrity of financial institutions.

Figure 1 shows that money laundering is dissected into three phases: placement, layering, and integration. It is worth noting that not all money-laundering transactions traverse through each of these phases.

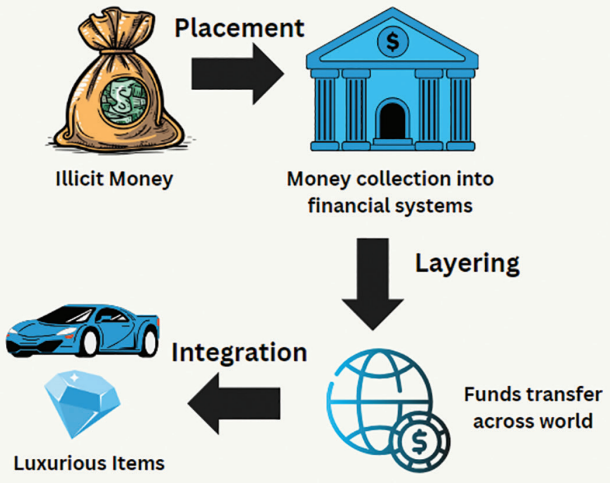

Figure 1: Money laundering: a three-stage process.

Nevertheless, delineating these distinct stages is imperative for comprehending this potentially intricate process. Traditional methodologies such as manual investigation, transaction monitoring, watchlist, and blacklist often result in high false-positive rates, rigidity, difficulty in handling big data, focus on individual transactions, and lack of adaptability [5,6].

Table 1 provides a comprehensive overview of significant fines imposed on companies and banks for AML violations. The fines are categorized by the name of the company or bank, the amount fined, the year of the penalty, and the reason for the fine.

Due to the continuous evolution of financial systems, the strategies applied to safeguard them must also evolve. Through empirical analysis, this research aims to discern the strengths and weaknesses of existing AML frameworks, exploiting the power of AI/ML algorithms to adapt and counteract emerging threats [7]. The findings of this study, rooted in real-world data, aspire to provide a robust foundation for the ongoing discourse on fortifying financial systems against the persistent and adaptive landscape of money laundering. Ultimately, the research seeks to contribute to the arsenal of tools available to financial institutions, regulatory bodies, and policymakers in their collective pursuit of securing the integrity of global financial systems.

This research paper addresses the critical gap by delving into the integration of powerful AI/ML algorithms such as gradient boosting regressor, least absolute shrinkage and selection operator (LASSO) regression, elastic net regressor, random forest regressor, and linear regression as potent tools for fortifying AML frameworks [8]. The paper aims to identify patterns and trends associated with money-laundering activities, considering temporal, geographical, and industry-specific variations. This involves analyzing data across different states, industries, and types of suspicious activities to uncover insights. Furthermore, the research seeks to enhance AML strategies by leveraging insights gained from the analysis of AI/ML algorithms. By identifying strengths and weaknesses, the study offers recommendations for optimizing AML strategies and informs policy decisions and regulatory frameworks in this domain. Additionally, the paper aims to facilitate the adoption of AI/ML technologies within the financial industry for AML purposes. Through showcasing tangible benefits and addressing potential challenges, the research aims to contribute to global financial security.

The paper is structured as follows. Section II provides a review of related literature; Section III

Table 1: The rising cost of AML shortcomings, urging stricter compliance across diverse sectors

Company name/bank	Fine	Year	Reason for fine
Binance Holdings Ltd (US)	$4 billion	2023	Breaches of Bank Secrecy Act, failure to register as money transmitter, violations of the International Emergency Economic Powers Act
Crown Resorts Ltd (Australia)	$450 million	2023	Past infractions of Australian AML regulations at casinos
Deutsche Bank (Germany)	$186 million	2023	Insufficient efforts to remedy money-laundering control and other weaknesses
Bank of Queensland (Australia)	$50 million (potential)	2023	Breaches of prudential norms and AML regulations
William Hill & Mr Green (UK)	£19.2 million	2023	Violations of AML and social responsibility regulations
Guaranty Trust Bank UK Ltd	£7.6 million	2023	Serious flaws in AML procedures and controls
ADM Investor Services International Ltd (UK)	£6.47 million	2023	Inadequate AML procedures and controls
In Touch Games Ltd (UK)	£6.1 million	2023	Failing to adequately handle money-laundering and social responsibility issues
HSBC (Mexico and Colombia)	$1.9 billion (£1.2 billion)	2023	Inadequate controls against money laundering
Credit Suisse Group (US)	$536 million	2009	Money-laundering allegations
Lloyds Banking Group PLC (UK)	$350 million	2009	Money-laundering allegations
ING Bank Group (the Netherlands)	$619 million	2012	Facilitating illegal movement of billions through the US banking system
Standard Bank PLC (UK)	$7.6 million	2014	Shortcomings in AML controls

AML, anti–money laundering.

Source: https://sanctionscanner.com/blog/the-five-biggest-money-laundering-scandals-317.

discusses about regulatory frameworks and compliance; Section IV discusses the proposed architecture; Section V outlines the proposed approach; Sections VI–VIII present the results, discussion, and future work, respectively. The paper concludes with Section IX, namely, the "Conclusion."

II. Related Work

Money laundering—the method of obscuring the origin of unlawfully acquired funds—poses a significant risk to global financial systems. Criminals engage in this process to legitimize their ill-gotten gains through a series of transactions, making it challenging for authorities to trace the illicit funds. The global financial landscape is increasingly plagued by the insidious threat of financial crimes, with money laundering standing out as one of the most pervasive and challenging illicit activities. Money laundering not only facilitates criminal enterprises but also undermines the integrity and stability of financial systems worldwide. Despite concerted efforts to combat this menace, traditional AML methods have needed to catch up with the evolving sophistication of illicit actors and their tactics. Considering these challenges, there is a growing recognition of the urgent need for innovative solutions to fortify AML frameworks and enhance the effectiveness of detection and prevention measures. One such promising avenue lies in the integration of AI and ML algorithms, which offer the capability to revolutionize the fight against money laundering by harnessing the power of data analytics and automation.

The research paper aims to explore the role of AI and ML in augmenting AML efforts within the financial sector. Drawing upon a comprehensive dataset

sourced from the US Treasury's FinCEN, this study seeks to unravel the complex patterns and anomalies inherent in money-laundering activities [9]. By leveraging advanced analytical techniques, including pattern recognition and anomaly detection, this research endeavors to uncover insights into the dynamics of money laundering and its interactions with contextual factors such as regulatory landscapes, industry sectors, and geographic locations. Through empirical analysis and interpretation of findings, this paper strives to provide valuable insights for AML practitioners, policymakers, and regulatory bodies, shedding light on potential vulnerabilities and strengths within current AML frameworks. Moreover, this paper critically examines the implications of integrating AI and ML technologies into AML practices, considering both the opportunities they present and the challenges they entail. By offering a nuanced understanding of the practical applicability of AI/ML in real-world AML scenarios, this research aims to contribute to ongoing efforts to heighten the resilience of financial institutions against the ever-evolving threat of money laundering.

The study was designed to leverage AI and ML algorithms to analyze patterns and anomalies in a unique dataset from FinCEN, focusing on various dimensions of money-laundering activities. The objective was to develop predictive models capable of identifying suspicious activities with high accuracy and to compare the effectiveness of different algorithms in this context. The study was initiated with the following steps:

- Data collection: Data were sourced and extracted from the FinCEN of the US Treasury, covering the period from 2014 to 2023.
- Data preprocessing: The dataset underwent extensive preprocessing, including normalization and encoding of categorical variables. This ensured that the data were clean and suitable for ML models.
- Algorithm selection: Various AI and ML algorithms were selected based on their proven capabilities in pattern recognition and anomaly detection. The algorithms included random forest classifier, elastic net regressor, LASSO regression, gradient boosting regressor, linear regression, multilayer perceptron (MLP) classifier, convolutional neural network (CNN), and *K*-nearest neighbor (KNN).
- Model training and validation: The data were split into training and validation sets. Each model was trained using the training set and validated using the validation set to ensure robust performance.
- Hyperparameter tuning: Extensive hyperparameter tuning was conducted using grid search and randomized search to optimize model performance.

The data were analyzed based on the following aspects:

- Performance metrics: The models were evaluated using various performance metrics such as accuracy, mean squared error (MSE), *R*-squared (R^2), precision, recall, and area under the receiver operating characteristic curve (ROC_AUC) scores.
- Comparative analysis: A comprehensive comparison of the models was conducted to identify the best-performing algorithms for different types of suspicious activities. This analysis highlighted the strengths and weaknesses of each model.
- Interpretation of results: The results were interpreted to understand the practical applicability of the models in real-world AML scenarios. Insights were derived regarding the effectiveness of AI/ML algorithms in detecting money-laundering patterns.

The key novel contributions of this research include a multifaceted approach that combines advanced AI/ML algorithms with extensive hyperparameter optimization techniques. Further, the study provides a comprehensive comparison of various AI and ML algorithms, such as random forest classifier, elastic net regressor, LASSO regression, gradient boosting regressor, linear regression, MLP classifier, CNN, and KNN, to evaluate their effectiveness in AML. The uniqueness of the FinCEN dataset adds significant value by providing rich, multifaceted insights crucial for developing more effective AML strategies. Extensive hyperparameter tuning through grid search and randomized search was conducted to optimize model performance, addressing a critical gap in current AML research. Furthermore, the empirical analysis and interpretation of findings offer valuable insights for AML practitioners, policymakers, and regulatory bodies, highlighting potential vulnerabilities and strengths within current AML frameworks. The evolution of AML measures has been a response to the growing complexity of financial crimes. Historical developments, such as the rise of transnational criminal organizations and the increased interconnectedness of global economies, have underscored the need for robust regulatory frameworks [10].

a. Global impact of money laundering

Over the years, the international community has recognized the imperative need to combat money

laundering, leading to the establishment of various AML regulations and frameworks. Regulatory bodies at the regional, national, and global levels have played a vital role in defining and enforcing compliance standards.

b. Rise of big data in finance

Moreover, the broader landscape of financial services has witnessed a technological revolution, with financial technology (FinTech) playing a central role. This evolution necessitates a concurrent innovation in AML strategies to address the dynamic nature of financial crimes and stay one step ahead of increasingly sophisticated criminal activities [11].

c. Emergence of AI and ML in AML

The emergence of AI and ML technologies has introduced a paradigm shift in AML strategies. These advanced technologies offer the promise of overcoming the limitations of traditional methods [12]. By delving into the interplay of technology, regulation, and financial crime, this study aims to contribute valuable insights to the ongoing efforts to fortify global financial systems against the pervasive threat of money laundering.

d. Traditional methods

The landscape of AML has witnessed a historical reliance on traditional methods, wherein rule-based systems, watchlists, and transaction monitoring form the backbone of detection strategies [13,14].

Table 2 discusses the traditional methods that are being applied for money-laundering-related activities.

d.i. Watchlists and blacklists

- Traditional approach: These lists contain known entities or activities associated with money laundering, and financial institutions use them to flag potentially suspicious transactions. Watchlists and blacklists are constructed and maintained by various government agencies and international organizations, such as the Financial Action Task Force (FATF), International Criminal Police Organization (Interpol), and national financial intelligence units (FIUs). The sources of information used to compile and update these lists include law enforcement records, intelligence reports, regulatory filings, and information shared by international partners. Regular updates are necessary to ensure that the lists reflect the latest intelligence on money-laundering activities and associated entities.
- Challenges: The main challenge is that these lists are static and do not capture new or evolving money-laundering techniques. Criminals can adapt their methods to avoid detection, making the reliance on static lists less effective over time.

d.ii. Transaction monitoring

- Traditional approach: This involves analyzing historical transaction data and applying predefined rules to detect potentially suspicious patterns or activities.
- Challenges: Transaction-monitoring systems often rely heavily on historical data, which can make them less effective at identifying emerging patterns or new money-laundering schemes. Additionally, the high volume of transactions and the manual

Table 2: Traditional AML methods struggle with dynamic schemes, outdated data, and manual burden

Traditional methods	Challenges
Watchlists and blacklists [15]	Limited effectiveness in identifying novel or evolving money-laundering schemes, reliance on static lists
Transaction monitoring [15]	Overreliance on historical data, potential to miss emerging patterns, and high manual review workload
CDD [16]	Difficulty in maintaining up-to-date customer profiles, potential for false negatives in risk assessments
Manual investigations [16]	Are labor-intensive, prone to human error, and may result in delays in identifying suspicious activities

AML, anti–money laundering; CDD, customer due diligence.

review required can lead to missed detections or delays in flagging suspicious activities.

d.iii. The CDD approach

- Traditional approach: CDD involves verifying the identity of customers, assessing their risk level, and monitoring their transactions for suspicious behavior based on predefined criteria. Financial institutions integrate CDD procedures into their customer-onboarding and transaction-monitoring workflows to ensure compliance with regulatory requirements and to mitigate risks associated with money laundering and other illicit activities. During the onboarding process, institutions collect and verify customer information, such as identification documents, financial background, and the purpose of the account. This information is used to create a risk profile for each customer.

Once customers are onboarded, their transactions are continuously monitored against their risk profile and predefined criteria to detect any unusual or suspicious activities. Advanced analytics and ML algorithms are often used to enhance the monitoring process, enabling institutions to identify patterns and anomalies indicative of potential money-laundering activities.

- Challenges: Maintaining up-to-date customer profiles and risk assessments is a significant challenge in traditional CDD methods. Outdated or incomplete information can lead to false negatives, whereby high-risk customers are not identified correctly. Additionally, ensuring that all relevant customer information is accurately captured and regularly updated requires substantial resources and coordination across various departments within the financial institution.

d.iv. Manual investigations

- Traditional approach: AML professionals conduct manual investigations to further analyze suspicious activities flagged by automated systems or through other means.
- Challenges: Manual investigations are labor-intensive, time-consuming, and prone to human error. The sheer volume of alerts generated by automated systems can overwhelm investigators, leading to delays in identifying and responding to suspicious activities.

e. Modern methods

In response to the faults of traditional methods, there has been a paradigm shift toward incorporating sophisticated methods, such as ML, predictive analytics, and network analysis, in the pursuit of more robust and adaptive AML strategies [17].

Table 3 outlines the advanced techniques used in AML and their associated challenges. It becomes evident that while these sophisticated approaches hold promise in bolstering AML efforts, they still face hurdles. These challenges encompass a spectrum ranging from data-related issues (such as availability, quality, and interpretability) to the dynamic nature of evolving threats posed by increasingly sophisticated criminals. Furthermore, the need for continuous improvement and adaptation to emerging tactics underscores the imperative for ongoing research and innovation in the

Table 3: Advanced techniques for AML: challenges in data labeling, interpretability, and evolving threats

ML algorithms/techniques	Challenges
ML algorithms [18]	Need for substantial labeled data, interpretability concerns, and potential biases in training data
Predictive analytics [19]	Dependence on accurate historical data, challenges in predicting novel or emerging techniques
NLP [20]	Handling diverse language nuances, extracting meaningful insights from vast textual data
Anomaly detection [21]	Balancing sensitivity and specificity, adapting to evolving tactics of sophisticated criminals
Big data analytics [22]	Ensuring scalability, data quality, and the need for robust infrastructure

AML, anti–money laundering; ML, machine learning; NLP, natural language processing.

field of AML. In the subsequent sections, we propose a framework for cohesively integrating advanced techniques to enhance the efficacy of AML strategies.

III. Regulatory Frameworks and Compliance

Regulatory compliance is a cornerstone of effective AML efforts, ensuring that financial institutions adhere to legal standards and guidelines designed to detect and prevent illicit financial activities. This section provides an overview of key AML regulations and discusses strategies used by organizations to maintain compliance while optimizing operational efficiency.

a. Overview of key AML regulations

Several key regulations and standards govern AML practices globally. These include the following:

- Bank Secrecy Act (BSA): Enacted in 1970, the BSA requires US financial institutions to assist government agencies in detecting and preventing money laundering. It mandates record-keeping and reporting certain transactions to authorities.
- Uniting and Strengthening America by Providing Appropriate Tools Required to Intercept and Obstruct Terrorism (USA PATRIOT) Act: Passed in 2001, the USA PATRIOT Act strengthens AML laws by requiring financial institutions to establish AML programs, verify the identities of customers, and report suspicious activities.
- FATF guidelines: FATF, an intergovernmental body, sets international standards to combat money laundering and terrorist financing. Its recommendations provide a framework for countries to establish robust AML/Combating the Financing of Terrorism (CFT) regimes.
- Other relevant international standards: Various countries have their own AML regulations that align with FATF guidelines, such as the European Union's Anti-Money Laundering Directives (AMLDs).

b. Compliance strategies in financial institutions

Financial institutions apply several strategies to comply with these regulations:

- Implementation of compliance measures: This involves developing comprehensive AML programs that include CDD, transaction monitoring, and reporting suspicious activities.
- Operational challenges and solutions: Institutions face challenges such as data quality issues, evolving regulatory requirements, and resource constraints. Solutions include investing in advanced technology, enhancing staff training, and collaborating with regulatory bodies.
- Use of technology (RegTech): Regulatory technology, or RegTech, leverages advanced technologies, such as ML, AI, and blockchain, to streamline compliance processes. These tools enhance the efficiency and effectiveness of AML programs by automating data collection, analysis, and reporting.

c. Balancing compliance and operational efficiency

Maintaining compliance while optimizing operational efficiency is critical. Strategies to achieve this balance include the following:

- Automated systems: Implementing automated transaction-monitoring systems to reduce manual efforts and improve accuracy.
- Risk-based approach: Prioritizing resources toward high-risk areas and activities, allowing for more effective allocation of compliance efforts.
- Continuous training: Providing ongoing training to staff, enabling them to stay updated on regulatory changes and best practices in AML compliance.

d. Blockchain technology for compliance and regulatory reporting in money service businesses (MSBs)

Blockchain technology offers significant potential to enhance compliance and regulatory reporting processes for MSBs while ensuring the privacy and security of sensitive financial data. The decentralized and immutable nature of blockchain can address many of the challenges faced by MSBs in adhering to regulatory requirements [23].

- Immutable record-keeping: Blockchain provides a tamper-proof ledger that ensures the integrity of transaction records. Every transaction recorded on a blockchain is immutable, meaning that it cannot be altered or deleted. This feature is crucial for compliance and regulatory reporting as it provides a transparent and verifiable audit trail. Regulators

can access historical transaction data with confidence in its authenticity, reducing the risk of fraud and ensuring accountability [23].

- Enhanced data privacy and security: Blockchain's cryptographic features protect sensitive financial data. Transactions on a blockchain are encrypted, making it extremely difficult for unauthorized parties to access or tamper with the data. Additionally, blockchain can apply advanced privacy-preserving techniques such as zero-knowledge proofs, which allow the validation of transaction details without revealing the actual data. This ensures that MSBs can maintain customer privacy while still complying with regulatory reporting requirements [23].
- Real-time monitoring and reporting: Blockchain enables real-time transaction monitoring, allowing MSBs to instantly report suspicious activities to regulators. Smart contracts, which are self-executing contracts with the terms directly written into code, can be programmed to automatically flag and report transactions that meet certain criteria for suspicious activity. This real-time capability enhances the effectiveness of AML efforts and ensures timely compliance with reporting obligations [23].
- Streamlined know your customer (KYC)/AML processes: Blockchain can streamline KYC and AML processes by providing a single, immutable record of customer identity and transaction history. Shared KYC platforms on a blockchain allow MSBs to verify customer identities more efficiently, reducing duplication of effort and improving the accuracy of customer data. Once a customer's identity is verified, it can be securely shared across the network, ensuring consistency and compliance with regulatory standards [23].

IV. Data

Figure 2 illustrates the data collection and processing pipeline. The initial step involves data collection, whereby information is gathered utilizing the customized FinCEN dataset for further processing. Following data collection, data cleaning and processing—entailing the removal of errors, data formatting, and transforming the data into a usable format—are carried out. Subsequently, the data are utilized for modeling purposes, wherein various modeling techniques, such as LASSO regression, gradient boosting regressor, random forest classifier, elastic net regressor,

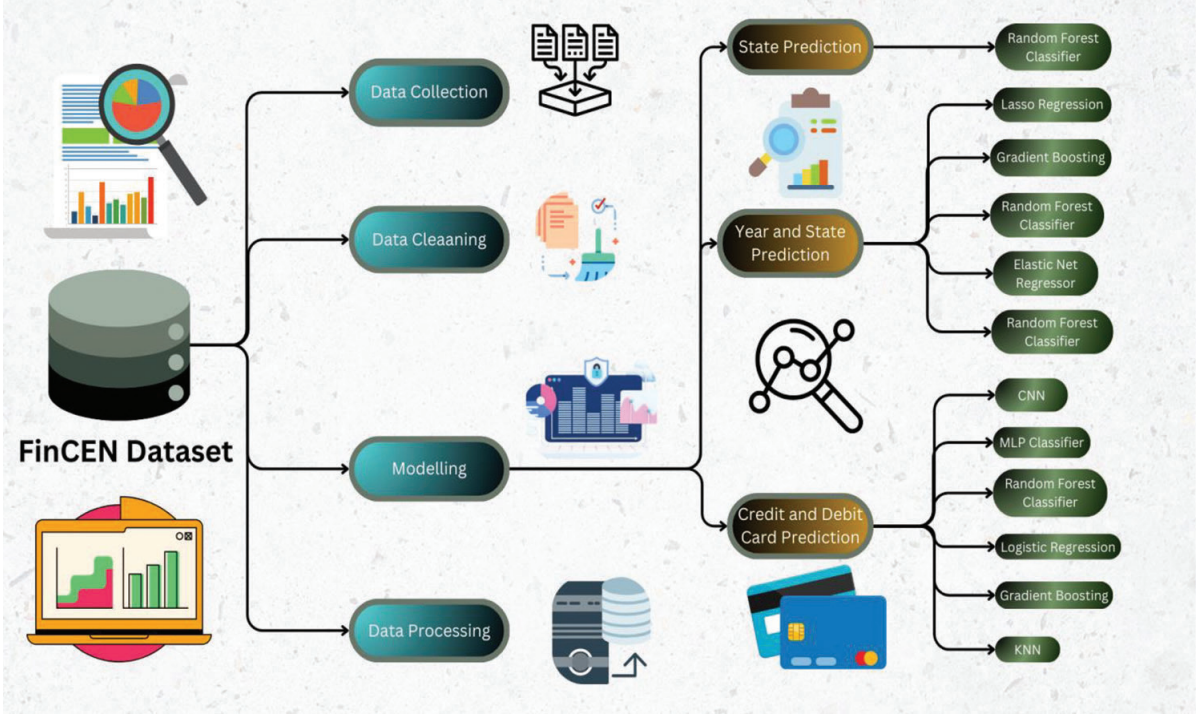

Figure 2: System flow architecture in AML modeling. AML, anti–money laundering; CNN, convolutional neural network; FinCEN, Financial Crimes Enforcement Network; KNN, K-nearest neighbor; MLP, multilayer perceptron.

MLP classifier, KNN, and CNN, are used to create predictive models.

a. Data collection

The FinCEN collects, analyzes, and disseminates financial transaction data to detect and investigate illicit activities such as money laundering and terrorist financing. FinCEN gathers data from various financial institutions and businesses, which are required by law to report suspicious transactions. These data are then analyzed to identify patterns, trends, and anomalies indicative of illegal activities. The insights generated are shared with law enforcement agencies and other relevant authorities to support their investigations and enforcement actions.

The types of data sources are as follows:

- Bank/cashier's checks, fund transfers, and foreign currency: Financial transactions that indicate money laundering or other illicit activities.
- Gaming instruments and government payments: Transactions in industries prone to money laundering.
- Money orders, personal/business checks, and traveler's checks: Common methods for transferring funds.
- US currency and other payment mechanisms: Tracks cash transactions and other forms of payments.

FinCEN's customized dataset contains various industry types and product types, including but not limited to the following:

- Product types: Bonds/notes, commercial mortgages, credit/debit cards, forex transactions, deposit amounts, futures/options, hedge funds.
- Data source: The data are collected from FinCEN's official reports, focusing on suspicious activities reported by MSBs across different states from 2014 to 2023. Each record in the dataset represents a specific incident of suspicious activity, covering key attributes such as transaction amounts, transaction types, and reporting institution details.

The customized FinCEN dataset contains different industry types, which include payment mechanisms such as bank/cashier's checks, fund transfers, foreign currency, gaming instruments, government payments, money orders, personal/business checks, traveler's checks, US currency, and others. Product types include bonds/notes, commercial mortgage, credit card, debit card, forex transactions, deposit amounts, future/options, hedge funds, and others. The data were collected from https://www.fincen.gov/reports/sar-stats. The dataset provides a snapshot of suspicious financial activities reported to the FinCEN of the US Treasury, focusing on MSBs across different states from the year 2014 to 2023. Each record in the dataset represents a specific incident of suspicious activity and includes the key attributes described in Section IV.b.

Privacy concerns and data protection regulations are critically managed to ensure the confidentiality and integrity of sensitive customer information. FinCEN uses stringent security measures, including encryption, access controls, and anonymization techniques, to protect the data from unauthorized access and breaches. Compliance with data protection laws, such as the General Data Protection Regulation (GDPR) and other relevant national regulations, is strictly enforced, ensuring that the handling and processing of data adhere to the highest standards of privacy and security.

The technologies applied are as follows:

- Encryption: Ensuring that data are encrypted both in transit and at rest to protect them from unauthorized access.
- Access controls: Implementing strict access controls to ensure that only authorized personnel can access sensitive data.
- Anonymization: Applying anonymization techniques to protect the identities of individuals involved in reported transactions.

To provide a more comprehensive view of potential money-laundering activities, it is essential to integrate data from multiple sources. Currently, our model primarily leverages transaction data. However, future iterations should incorporate additional data sources such as customer profiles and external data (e.g., social media activity, news reports). This integration can offer a richer context, enhancing the model's ability to detect complex laundering schemes. By incorporating these diverse datasets, we can develop more robust and accurate models, ultimately improving our ability to identify and prevent money laundering.

b. Data description

Section IV.b describes the dataset utilized in the study, comprising the following variables:

- Year: This is the year in which the suspicious activity was reported, consisting of data from 2014 to 2023.

- State: This is the geographical location of the 61 states that reported suspicious activity, indicating the specific state within the United States.
- Industry: This specifies the nature of the business involved in the suspicious activity, with a focus on MSBs.
- Suspicious activity: This describes the type or nature of the suspicious activity reported. The dataset includes all types of money laundering such as exchanging small bills for large bills or vice versa, funnel accounts, suspicion concerning the physical condition of funds, suspicion concerning the source of funds, suspicion of designation of beneficiaries, suspicious electronic fund transfers (EFTs)/ wire transfers, exchange of currencies, receipt of government payments/benefits, use of multiple accounts, use of noncash monetary instruments, use of third-party transactions, trade-based money laundering, transaction out of pattern for customers, and other money-laundering activities.
- Product type: This variable indicates the financial product associated with the suspicious activity. The dataset contains credit card and debit cards as product types.
- Instrument type/payment mechanism: It specifies the financial instrument used in the reported activity; the dataset includes fund transfer as the instrument type.
- Regulators: This variable identifies the regulatory authority overseeing the reported activity, including CFTC, Federal Deposit Insurance Corporation (FDIC), Federal Housing Finance Agency (FHFA), Federal Reserve Board (FRB), Internal Revenue Service (IRS), National Credit Union Administration (NCUA), Office of the Comptroller of the Currency (OCC), Securities and Exchange Commission (SEC), and others.
- Count: This represents the count or frequency of occurrences of the reported suspicious activity, providing a quantitative measure.

This dataset offers insights into the geographical distribution, temporal patterns, and nature of suspicious financial activities within the MSB sector. It serves as a worthwhile resource to analyze trends, identify risk factors, and contribute to broader efforts in combating financial crimes, particularly in the context of money laundering.

c. Data processing

The raw data collected from the FinCEN dataset contain features such as "Year" (which is converted to the "Date–Time" format) and "extracted Year." For data preprocessing, the categorical features such as "State," "Industry," "Suspicious activity," "Regulator," "Product type," and "Instrument type" are label encoded, specifically one-hot encoding for categorical variables, ensuring compatibility with the RandomForestClassifier and RandomForestRegressor models [23].

d. Data analysis

Figure 3 presents a bar graph titled "Total suspicious activity count by product," illustrating the counts of suspicious activities associated with three different products: credit card, debit card, and Item1. The horizontal bar graph has orange bars representing the count of suspicious activities for three distinct products. The *x*-axis is labeled "Product" and has three categories: "Credit Card," "Debit Card," and "Total." The *y*-axis is labeled "Total count of suspicious activities" and ranges from 0 to 1.4. The "Credit Card" category has a significantly lower count compared to the other two categories. "Debit Card" and "Item1" have similar high counts of suspicious activities, with "Item1" having slightly more. The graph suggests that some types of suspicious activities are more prevalent than others and that the source of funds for these activities is often questionable.

Figure 4 is a line graph depicting the monthly trend of suspicious activity count over an unspecified period of years. The graph is titled "Monthly Trend of Suspicious Activity Count." The *x*-axis represents the years, although specific years are not labeled. The *y*-axis represents the total count of suspicious activities, ranging from 0 to 800,000. There is an orange line with data points marked by circles representing the trend in suspicious activity counts. Initially, there is a slow and steady increase in the count. Toward the end of the timeline, there is a sharp increase in suspicious activities, as indicated by the steep incline of the line. The graph suggests that some types of suspicious activities are more common than others and that the source of funds for these activities is often questionable.

Figure 5 presents a pie chart depicting the distribution of various suspicious activities. Different colors represent different types of suspicious activities, each labeled with a percentage indicating its proportion. The largest segment, colored in green and occupying 37.4%, represents "Suspicion concerning the source of funds." Another significant segment, colored blue and taking up 26.0%, is labeled "Exchanges/transfers/large transactions." Other

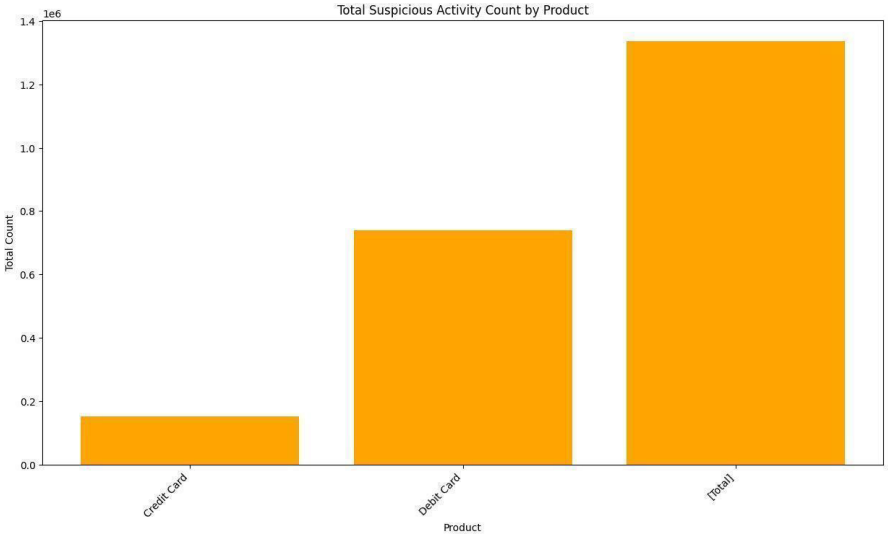

Figure 3: Illustration of the "differentiating suspicious activity counts," showcasing varying totals of suspicious activities across "Credit Card, Debit Card, and Total" categories. This variation suggests differing risk levels and potential financial irregularities associated with each category.

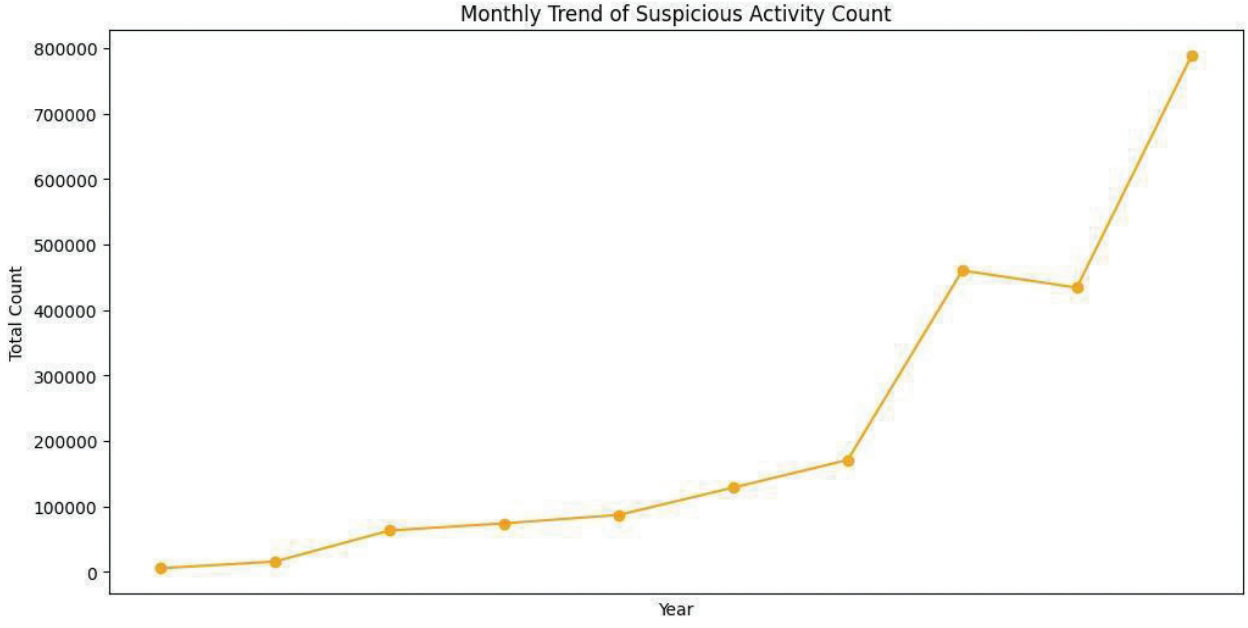

Figure 4: Tracking suspicious activity trends: figure shows a gradual rise, followed by a sharp increase in monthly suspicious activity counts over the years.

segments include "Suspicious use of multiple accounts," "Transaction out of pattern for customer(s)," and several more, each with varying percentages. Each segment's label includes detailed descriptions or examples of the type of suspicious activity that it represents. The pie chart suggests that some types of suspicious activities are more prevalent than others and that the source of funds for these activities is often questionable.

Figure 6 depicts the distribution of suspicious activity reports (SARs) filed with the FinCEN across various states in the United States. Each bar represents the total number of reported suspicious activities related to potential money laundering within that

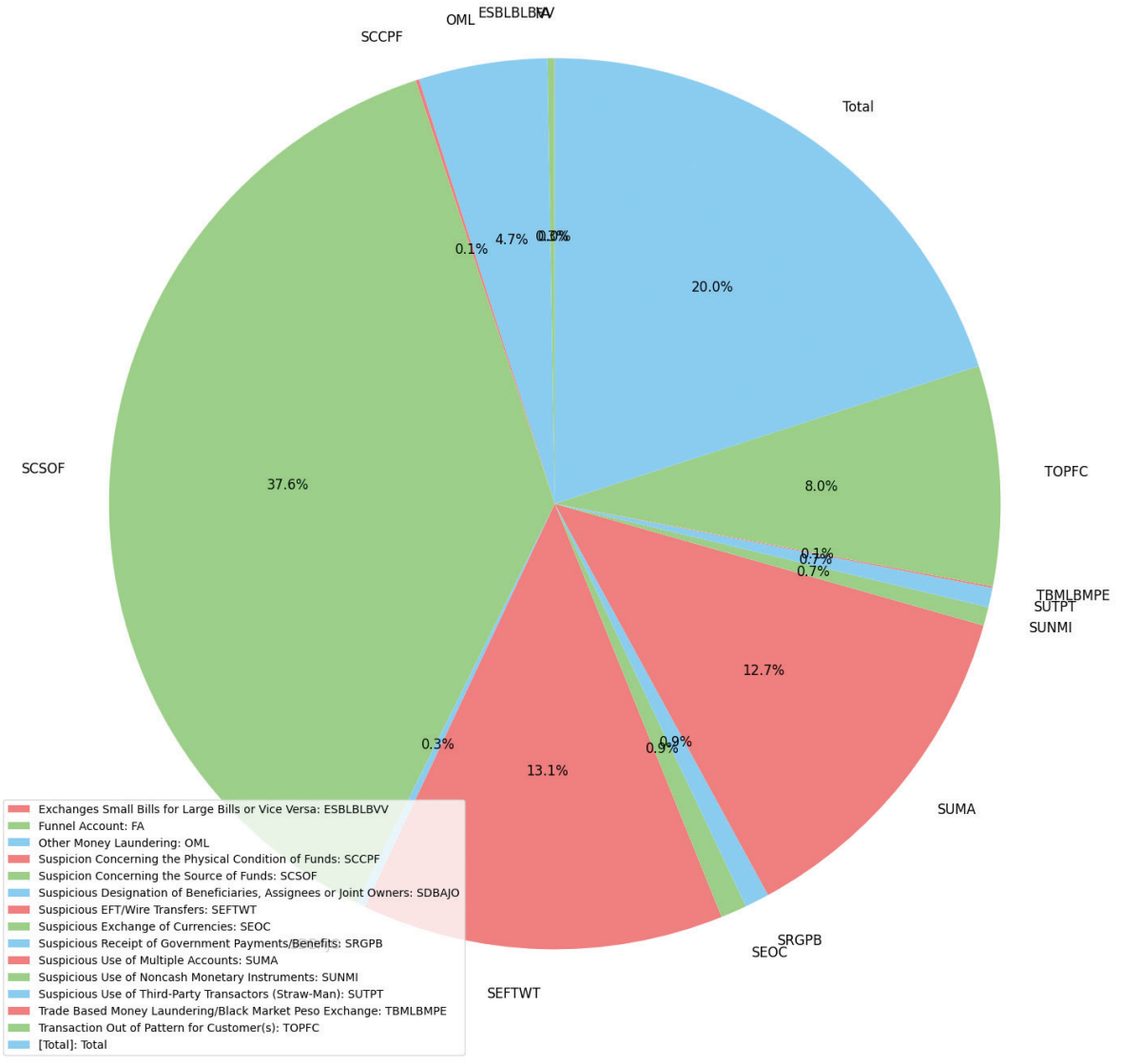

Figure 5: The "spectrum of suspicious activities," showing—in a pie chart format—a breakdown that highlights prevalent concerns in financial transactions, including "Source of funds" and "Exchanges/transfers," among others.

state. California, Colorado, Texas, and Florida have the highest number of SARs filed, with >100,000 each. This is due to a number of factors, including the large populations of these states, their major financial centers, and the fact that they are coastal states, which are often seen as more vulnerable to money laundering. Alaska, Wyoming, and Vermont have the lowest number of SARs filed, with <10,000 each. This is due to their smaller populations and more rural economies. There is a general trend of higher SAR filings in the more populous states in the east and west of the country and lower filings in the less populous states in the center of the country.

Figure 7 displays the number of SARs filed with various regulatory agencies in the United States. The *y*-axis represents the count of suspicious activities, while the *x*-axis lists the corresponding regulators. The IRS received the highest number of SARs, >100,000. This is due to the agency's broad oversight of financial transactions and its focus on tax-related money laundering. The FRB reported "Not applicable," indicating that it does not handle SARs directly. The remaining regulators, including the FDIC, NCUA, OCC, and SEC, had lower counts of suspicious activities compared to the IRS [24].

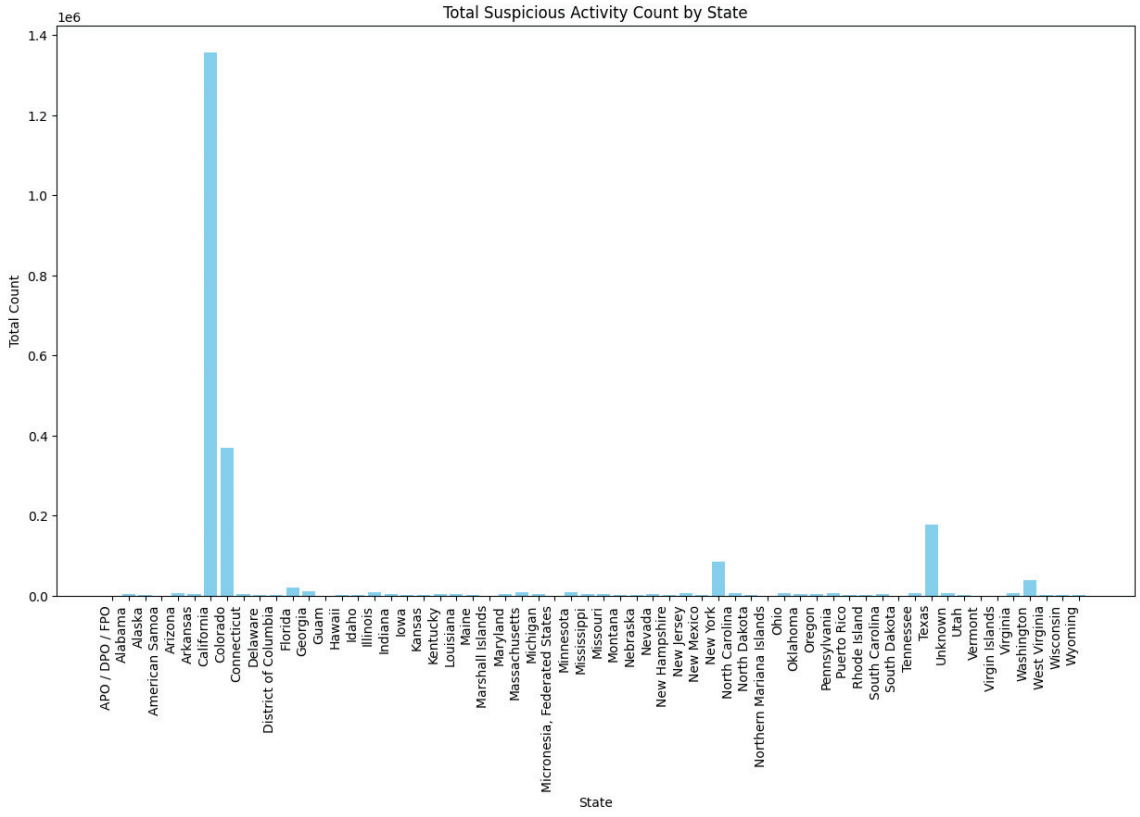

Figure 6: Regional disparities in SAR filings are shown, depicting high activity in populous coastal states versus lower reports in rural areas, signaling potential money-laundering vulnerabilities. SAR, suspicious activity report.

V. Proposed Approach

The selection of ML algorithms for this research was intricately tied to the features present in the dataset and the specific objectives of the study.

a. Data preprocessing and feature engineering

Data preprocessing involves cleaning and preparing the dataset for analysis. This includes handling missing values and encoding categorical variables. Feature engineering aims to extract relevant information from the dataset to improve prediction accuracy.

b. Model selection and optimization

b.i. Prediction of state

- Random forest classifier: It is a robust and versatile ML algorithm commonly used for classification tasks, making it well-suited for predicting the "State" variable in financial transactions. It belongs to the ensemble learning family of algorithms, wherein numerous decision trees are trained on distinct subsets of the data, and their predictions are compiled to produce the final output [2,4].
- Accuracy and robustness: The random forest classifier was chosen for its high accuracy and robustness, even when dealing with complex datasets with high dimensionality and its ability to prevent overfitting through ensemble learning. In the context of predicting the geographical origin ("State") of financial transactions, accuracy is crucial to ensure that transactions are correctly attributed to their respective states.
- Handling imbalanced data: In financial transactions, the occurrence of fraudulent transactions is often rare compared to legitimate transactions, leading to imbalanced datasets. The random forest classifier is robust to class imbalance and can effectively handle skewed data distributions. This approach increases the likelihood that minority class instances (fraudulent transactions) are included in the training subsets. Furthermore, the algorithm aggregates the predictions from all

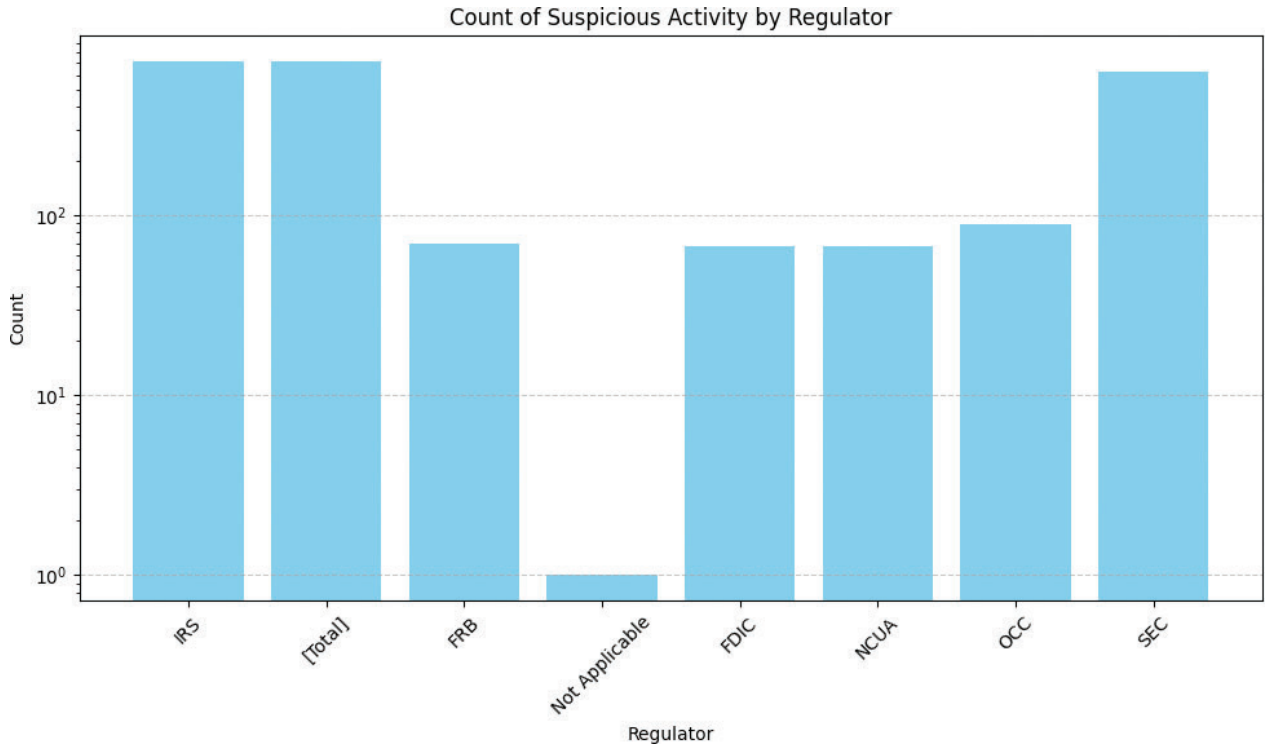

Figure 7: IRS leads in SAR Filings: figure reveals IRS as the primary recipient of SARs, reflecting its extensive oversight in financial transactions. FDIC, Federal Deposit Insurance Corporation; FRB, Federal Reserve Board; IRS, Internal Revenue Service; NCUA, National Credit Union Administration; OCC, Office of the Comptroller of the Currency; SARs, suspicious activity reports; SEC, Securities and Exchange Commission.

trees, which helps to balance the influence of the minority class, thereby improving the detection of fraudulent transactions and reducing the bias toward the majority class. This ensemble method enhances the overall performance and reliability of the classifier in identifying rare fraudulent activities amid a large volume of legitimate transactions [24].

b.ii. Prediction of year and state simultaneously

- Elastic net regressor: Elastic net is a regularization technique that combines the penalties of both L1 (LASSO) and L2 (ridge) regularization methods. It is majorly useful when dealing with high-dimensional datasets and helps in feature selection by shrinking coefficients and encouraging sparsity. The elastic net regressor aims to reduce the total sum of squared differences between the observed and predicted values while also penalizing the model for the magnitude of the coefficients. This metric, known as the residual sum of squares (RSS), measures the discrepancies between the actual data points and the values predicted by the model. Minimizing the RSS is crucial because it directly correlates with the model's accuracy. The smaller the RSS, the closer the predicted values are to the actual values, indicating a better fit of the model to the data. By reducing the total sum of squared differences, the model aims to improve its predictive performance, ensuring that the predictions are as accurate as possible based on the given data [25].
- LASSO regression: LASSO regression, also known as L1 regularization, adds the absolute value of the coefficients as a penalty term to the loss function. This encourages sparsity in the coefficient estimates and effectively performs feature selection by shrinking some coefficients to zero. LASSO regression is especially useful when dealing with high-dimensional datasets with many irrelevant or redundant features.
- Random forest regressor: The random forest regressor is a powerful ensemble learning algorithm used for supervised regression tasks. It is

directed by constructing multiple decision trees during training, with each tree being trained on a random subset of the training data and features. Through aggregation of predictions from individual trees, typically by averaging, the final output is obtained. The random forest regressor is an ensemble learning method that mitigates overfitting and enhances generalization through its unique approach. It builds multiple decision trees using bootstrap sampling (where each tree is trained on a different subset of the data) and random feature selection (where a random subset of features is considered at each split). This introduces diversity among the trees, reducing the risk of overfitting to any specific data subset or feature set. By aggregating the predictions from multiple trees, the random forest algorithm averages out individual-tree errors, resulting in more robust and accurate predictions [24]. This ensemble approach lowers model variance, making it less sensitive to noise and fluctuations in the training data, thus improving performance on unseen data. In our analysis, the random forest regressor effectively predicted continuous outcomes, demonstrating strong accuracy and stability. Its ability to handle diverse data subsets and features ensured reliable predictions and reduced overfitting compared to single decision trees or less-robust algorithms. Its capability to capture complex relationships in the data while maintaining interpretability makes it a popular choice for predicting continuous variables in fields such as finance, health care, and environmental science.

- Gradient boosting regressor: Gradient boosting is an ensemble learning approach that constructs a sequence of weak learners (typically decision trees) in a forward stage-wise manner. Each subsequent model focuses on minimizing the errors made by the previous models. The gradient boosting regressor optimizes a loss function using gradient descent and updates the model by fitting it to the negative gradient of the loss function. Initially, the model starts with an initial prediction, often the mean of the target values. Residuals (errors) are then calculated as the difference between the actual values and the predicted values from the current ensemble. A new decision tree is fitted to these residuals, effectively learning to predict the negative gradient of the loss function with respect to the current model's predictions. The predictions are updated by adding a fraction, controlled by the learning rate, of the new model's predictions to the existing ensemble's predictions. This iterative process ensures that each new model reduces the overall loss by focusing on the errors of the current model, thereby optimizing the model's performance through gradient descent and improving accuracy incrementally [11].
- Linear regression: Linear regression is an uncomplicated and easily understood algorithm used for regression analysis. It establishes a linear relationship between predictors (independent variable) and a target (dependent variable). It diminishes the sum of the squared differences between the observed and the predicted values, with coefficients indicating the gradient of the linear association between each predictor and outcome variable.

b.iii. Prediction of transactions that occurred through credit card and debit card

- MLP classifier (deep learning [DL]): The MLP classifier is a type of feedforward neural network that consists of multiple layers of nodes (neurons), comprising an input layer, several hidden layers, and an output layer [25]. Each node within a layer is linked to every node in the subsequent layer, and each connection has an associated weight. During training, the network learns to adjust the weights to minimize the error between the predicted and the actual outputs using optimization techniques such as gradient descent [25]. The MLP classifier was applied to learn intricate patterns in the dataset, which traditional linear models do not capture. The efficacy of the MLP classifier is significantly enhanced by the use of nonlinear activation functions such as the rectified linear unit (ReLU), sigmoid, or tanh. These activation functions allow the network to capture and model complex, nonlinear relationships between the input features and the target variable. This helps in capturing interactions and dependencies that are not immediately apparent in the raw data. Unlike linear models, which can only represent data through straight-line boundaries, nonlinear activation functions enable MLPs to learn and approximate intricate decision boundaries [26].
- CNN (DL): CNNs are DL neural network models particularly suited for processing structured grid-like data, such as images. They consist of various layers, such as convolutional layers, pooling layers, and fully connected layers. These are trained to automatically learn hierarchical patterns and features from the data. In our analysis, CNNs were applied to extract relevant features from the credit card and debit card transaction data, leveraging their

ability to capture spatial and temporal dependencies [27]. By learning hierarchical representations, CNNs can effectively discern intricate patterns in the data, leading to enhanced classification performance compared to traditional ML algorithms. The dataset includes various types of suspicious activities, such as exchanging small bills for large bills, use of multiple accounts, and trade-based money laundering. CNNs can effectively identify these patterns by learning from the data spanning from 2014 to 2023 across 61 states. The ability to process the temporal dimension of the data allows CNNs to recognize anomalies over time, such as unusual spikes in fund transfers.

- Random forest classifier: Random forest is an ensemble learning algorithm that works by creating multiple decision trees during training and introducing the class (distribution) model or mean prediction (regression) of each tree [25,28,29]. Each tree in the forest is made using a different set of features and a random subset of the training set; this helps reduce overfitting and improve generalization. GridSearchCV is a hyperparameter optimization technique in which a grid of hyperparameters is defined, and the best combination is selected based on competition. Its detailed description follows. In our implementation, the random forest classifier is trained using the grid search CV technique to optimize its hyperparameters, which include the maximum depth of each tree and the number of trees. This combination has proven useful in capturing the relationship between features and target variables to accurately predict credit and debit card transactions.
- Logistic regression: Logistic regression is a popular linear classification algorithm that models the probability of a binary outcome using the logistic function. It estimates the coefficients of the input features to make predictions, and regularization techniques such as L1 and L2 regularization are often used to prevent overfitting [20]. In our study, logistic regression was utilized to classify credit card and debit card transactions based on their features. Despite its simplicity, logistic regression can offer interpretable results and is computationally efficient, making it a suitable choice for binary classification tasks with interpretable coefficients.
- Gradient boosting: Gradient boosting is a collective learning method that amalgamates several basic predictors to form a stronger predictive model, typically decision trees. It sequentially trains new models to correct the errors of the previous models, with each new model focusing on the residuals of the previous ones [11]. In our experiments, gradient boosting was used to predict credit card and debit card transactions by iteratively improving the model's predictive performance. By focusing on the mistakes made by earlier models, gradient boosting can enhance the overall accuracy and robustness of the predictions, making it a powerful tool for classification tasks.
- KNN: KNN is a simple yet effective nonparametric algorithm used for classification tasks. It calculates the distance between the input data point and all other points in the dataset to determine the KNNs [28]. The class label is then assigned based on the majority class among the nearest neighbors. KNN was utilized in our analysis to classify credit card and debit card transactions based on the similarity of their features to neighboring instances. KNN is known to suffer from computational inefficiency, particularly when dealing with large datasets. This inefficiency arises because KNN requires calculating the distance between the query point and every point in the dataset for each prediction, leading to a significant increase in computational cost as the dataset grows [28]. The high-dimensional space often involved in such datasets further exacerbates this issue, as calculating distances in high dimensions is computationally intensive and can slow down the performance considerably.
- Deep neural networks (DNNs): In addition to traditional ML models, we explored the use of DL techniques, including DNNs, for predicting credit card and debit card transactions [30]. However, despite the potential of DL models to learn hierarchical representations of data, we encountered challenges that could have improved their effectiveness in this particular task. DNNs are extremely flexible models with a large number of parameters, which makes them prone to overfitting, especially when trained on limited data. Complexity and interpretability: DL models are inherently complex, making them difficult to interpret and understand the underlying decision-making process. In domains such as financial transactions and money-laundering detection, interpretability is crucial for regulatory compliance and risk assessment. The black-box nature of DL models hinders their adoption in applications wherein transparency and interpretability are crucial. One significant reason is the tendency of DL models to overfit, especially with limited or noisy data, leading to poor generalization on new, unseen data. Overfitting makes it difficult for stakeholders to trust the model's predictions. Additionally, the complex and opaque decision-making process of DL models prevents users from understanding how specific inputs are

transformed into outputs, raising concerns about bias and errors [11].

As mentioned herein, a diverse set of ML and DL algorithms are considered for predicting year and state variables, as well as credit card and debit card transactions. This includes regression algorithms such as elastic net regressor, LASSO regression, and gradient boosting, as well as classification algorithms such as random forest, logistic regression, MLP, CNN, and KNN.

c. Hyperparameter tuning

To enhance the predictive accuracy of ML algorithms for financial variables such as year, state, or transaction types, model hyperparameters are generally optimized using advanced techniques such as grid search or randomized search. This optimization process is pivotal in fine-tuning the algorithms to achieve optimal performance, ensuring robustness and reliability in predicting complex financial patterns. Grid search meticulously explores a predefined grid of hyperparameter values, systematically evaluating each combination through cross-validation to select the configuration yielding the highest performance metric [31]. Conversely, randomized search efficiently samples hyperparameter values from predefined distributions, making it particularly effective for large search spaces. Within the context of this research paper, leveraging grid search or randomized search ensures that models such as random forest or gradient boosting are meticulously fine-tuned, resulting in heightened predictive accuracy for financial variables. This systematic approach not only enhances the models' ability to capture intricate financial patterns but also contributes to their robustness and generalization capability, rendering them invaluable tools for real-world applications in financial crime detection and prevention.

d. Cross-validation and evaluation parameters

Rigorous cross-validation procedures were conducted to assess the generalization performance of the proposed models. This involved partitioning the dataset into training and test sets multiple times to ensure reliable performance estimation.

d.i. The MSE approach

The MSE calculates the median of the squared differences between the actual values and the predictions made by the model as shown in Eq. (1). It provides a measure of the median squared deviation of the predicted values from the true values. A lower MSE indicates that the model's predictions are closer to the actual values, suggesting better overall performance in terms of accuracy [32].

$$MSE = 1/n\sum\left(Predicted\,Rate - Actual\,Rate\right)^2$$

(range of summation varies from $i = 1$ to $i = n$) (1)

d.ii. The R^2 approach

The coefficient of determination R^2 measures the ratio of the variable in the sample (objective) that is explained by individual variables (characteristics). The values of R^2 range from "0" to "1." Here, "1" means that the model explains the variance in the data perfectly, while "0" means that the model explains no variance. The more optimized the R^2 value, the better the model fits the data and explains more of the variance. This indicates the effectiveness of the prediction [33].

d.iii. Mean absolute error (MAE)

The MAE calculates the average difference between the prediction and the actual result. It provides a measure of the average size of the error in the estimate. Compared to the MSE, the MAE is less sensitive to outliers because it has no error. It, therefore, provides a direct description of the performance of the model. A lower MAE means that the model's prediction is closer to the true value, indicating higher accuracy [32–34].

d.iv. Precision

Precision indicates the ratio of true positive predictions amid all positive predictions made by the model. It indicates the accuracy of the positive predictions made by the model. It is calculated as the ratio of true positives to the total sum of true positives and false positives, as mentioned in Eq. (2) [35].

$$Precision = \frac{True\,positives}{True\,positives + False\,positives} \quad (2)$$

A higher value for precision implies that the model makes fewer false-positive predictions, suggesting higher reliability in identifying true positives.

d.v. Recall

Recall, also known as the true positive value, measures the percentage of true positive predictions that the model correctly identifies among all positive cases. Recall is calculated as the ratio of true positives to the total number of true positives and false negatives, as stated in Eq. (3) [35].

$$Recall = \frac{True\ positives}{True\ positives + False\ negatives} \tag{3}$$

The higher the recall rate, the greater is the proportion of false positives captured by the standard.

d.vi. Accuracy

Accuracy is one of the most straightforward evaluation metrics and represents the ratio of correctly predicted instances to the total number of instances in the dataset. Mathematically, accuracy is calculated as the sum of true positives and true negatives divided by the sum of true positives, true negatives, false positives, and false negatives, as stated in Eq. (4) [35]:

$$Accuracy = \frac{True\ positives + True\ negatives}{True\ positives + True\ negatives + False\ negatives + False\ positives} \tag{4}$$

Accuracy provides a general overview of the model's performance and is especially useful when classes are balanced. However, there are better metrics to use when dealing with imbalanced datasets, as it can be misleading in such cases.

d.vii. F1 score

The F1 score is a metric that fuses recall and precision into a single value. It provides a balance between these two metrics and is particularly useful when classes are imbalanced [35]. Mathematically, the F1 score is the harmonic mean of precision and recall, as mentioned in Eq. (5):

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{5}$$

The F1 score is measured on a scale from "0" to "1," with a greater score reflecting the superior performance of the model. It penalizes models with imbalanced precision and recall values, making it suitable for evaluating models in situations where false positives and false negatives have different costs.

d.viii. Area under the receiver operating characteristic curve (ROC_AUC)

The ROC curve graphically depicts the trade-off between the true-positive and the false-positive rates across various thresholds. The ROC_AUC quantifies the overall performance of the model across all possible classification thresholds [36]. It depicts the probability that the model will rank a positive instance chosen at random higher than a chosen negative instance chosen at random. The ROC_AUC curve is particularly useful for evaluating binary classification models, especially when the classes are disproportionate or the cost of false positives and false negatives varies. Higher values of ROC_AUC indicate better model discrimination between positive and negative instances, making it a popular metric for assessing the performance of a classification model.

These evaluation parameters provide valuable insights into the performance of ML models for predicting financial variables. They help assess the accuracy, reliability, and effectiveness of the models in capturing patterns and relationships in the financial data, thereby informing decision-making and facilitating improvements in model performance.

By following the aforementioned comprehensive approach, we aim to develop robust and accurate prediction models for year and state variables, as well as credit card and debit card transactions, leveraging advanced ML and DL techniques to address complex prediction tasks in the financial domain.

VI. Results

a. State prediction

The random forest classifier was used for state prediction, resulting in an impressive performance summary. The average accuracy across all states was an exceptional 99.99%, showcasing the model's robustness and effectiveness in classifying states accurately. This high level of accuracy indicates the reliability of the random forest classifier in distinguishing between different states based on the provided features or variables.

Furthermore, the average confusion matrix reflects the model's ability to classify instances correctly. A confusion matrix serves as a valuable instrument for

assessing a classifier's effectiveness, detailing the count of correct identifications (true positives), correct rejections (true negatives), mistaken identifications (false positives), and missed detections (false negatives) for each category. In this case, the average confusion matrix would provide insights into how well the random forest classifier performed in terms of correctly identifying instances belonging to each state and minimizing misclassifications.

Table 4 presents the confusion matrix depicting the predicted versus actual classifications of financial transactions using the random forest classifier. The confusion matrix clearly shows the random forest classifier's ability to accurately predict both nonfraudulent and fraudulent transactions. Out of 3,811 nonfraudulent transactions, the classifier correctly predicted 3,689, resulting in a high true-negative rate. Similarly, out of 185 fraudulent transactions, the classifier correctly predicted 123, indicating a considerable true-positive rate. The small number of false positives ($n = 62$) and false negatives ($n = 62$) further demonstrates the model's robustness in distinguishing between fraudulent and nonfraudulent transactions.

This strong performance underscores the reliability of using the random forest classifier for predicting the geographical origin of financial transactions based on the "State" variable. With an exceptional average accuracy of 99.99% across all states, the model proves its effectiveness in accurately determining the state associated with each financial transaction. Such precision and reliability are crucial in financial analytics and fraud detection, highlighting the practical value of the random forest classifier in real-world applications. The model's high accuracy is evident from the large number of true positives ($n = 123$) and true negatives ($n = 3{,}689$) in the confusion matrix. The relatively low number of false positives ($n = 62$) indicates the model's ability to correctly identify nonfraudulent transactions. Similarly, the low count of false negatives ($n = 62$) highlights the model's effectiveness in recognizing fraudulent transactions.

Table 4: Confusion matrix illustrating the predicted versus actual classifications of financial transactions

	Predicted nonfraud	Predicted fraud
Actual nonfraud	3,689	62
Actual fraud	62	123

b. Prediction of year and state simultaneously

In evaluating different regression algorithms for their predictive performance, several key observations emerge (Table 5). First, the elastic net regressor stood out with the highest MSE of 3.25 and the highest MAE of 1.085. These metrics suggest that the elastic net regressor exhibited relatively poorer performance compared to other algorithms, indicating a greater degree of error in its predictions. The higher error rates of the elastic net regressor result from its combination of L1 (LASSO) and L2 (ridge) regularizations, which, while useful for handling multicollinearity and feature selection, do not always align well with the underlying data patterns and leads to underfitting when the true relationship between the features and the target variable is complex and is not captured adequately by a linear combination of features, compared to more specialized or sophisticated algorithms such as gradient boosting or random forest. This leads to the elastic net model producing a higher number of prediction errors. On the other hand, the LASSO regression algorithm demonstrated the lowest MAE of 0.84, signifying that, on average, it made predictions closest to the actual values. This suggests a higher level of accuracy and precision in its predictions compared to the other algorithms considered. Interestingly, the random forest algorithm excelled in terms of MSE, boasting a value of 2.5, which signifies the lowest error among all the algorithms evaluated. Furthermore, its second-lowest MAE of 0.8 reinforces its effectiveness in making accurate predictions,

Table 5: Performance metrics of different models for year and state

Algorithms	MSE	R^2	MAE
Elastic net regressor	3.25	0.10	1.085
LASSO regression	3.28	0.10	0.840
Random forest	2.50	0.60	0.800
Gradient boosting regressor	2.90	0.24	1.010
Linear regression	3.20	0.50	1.060

LASSO, least absolute shrinkage and selection operator; MAE, mean absolute error; MSE, mean squared error.

showcasing its robustness and reliability in regression tasks.

The gradient boosting regressor also showcased competitive performance with an MSE of 2.9 and the highest R^2 value among all algorithms at 0.24. This R^2 value indicates that the gradient boosting regressor can explain approximately 24% of the variance in the data, highlighting its ability to capture and model underlying patterns effectively.

Finally, the linear regression algorithm yielded an R^2 value of 0.5, suggesting that it can explain around 50% of the variance in the data. While this indicates a moderate level of explanatory power, it falls behind the other algorithms in terms of predictive accuracy and performance, as reflected in its higher errors compared to algorithms such as LASSO regression and random forest.

c. Credit card prediction

Table 6 indicates the performance of various classification algorithms and sheds light on their strengths and weaknesses in handling different types of data and tasks. MLP neural networks excel in capturing complex nonlinear relationships within data. However, the observation of lower precision compared to CNNs suggests that the MLP model might have misclassified some instances, indicating a potential limitation in handling specific types of data or patterns. On the other hand, CNNs are renowned for their prowess in image-processing tasks but can also be adapted to sequential data such as time series. Despite their versatility, the noted low recall in this context indicates that the CNN model struggled to identify true positives effectively, which could be attributed to the complexity of the data or model architecture.

In contrast, the random forest classifier emerges as a robust and versatile model capable of handling complex datasets with high dimensionality, making it particularly effective for tasks such as credit card classification. Its ability to both handle various types of features and capture intricate relationships contributes to its reliability and performance in classification tasks. Logistic regression, although a simpler algorithm, remains a powerful tool for binary classification tasks. Its linear nature makes it interpretable and efficient, especially when the relationship between features and the target variable is approximately linear. This simplicity often translates to ease of implementation and understanding while maintaining competitive performance in relevant contexts. Similarly, gradient boosting algorithms stand out for their iterative improvement of weak learners, focusing on rectifying mistakes made by previous models. This approach leads to high predictive accuracy, making gradient boosting a preferred choice in situations where accurate predictions are crucial. Lastly, KNN demonstrates notable recall, indicating its capability to capture a larger proportion of true positives. This characteristic makes KNN suitable for scenarios in which correctly identifying positive instances is of utmost importance, even though KNN comes with computational costs associated with its proximity-based approach.

d. Debit card prediction

Based on the findings from Table 7, the performance of various ML models in debit card classification can be summarized as follows. MLP, a type of DL model, showed significant promise with high precision and recall rates. The multilayer structure of MLPs enables them to capture intricate patterns in the data, making them particularly effective for this task. Conversely, CNNs, another DL approach, performed reasonably well but did not fully leverage the sequential nature of the data compared to MLPs. Random forest classifier,

Table 6: Performance metrics of different models for credit card classification

Model	Accuracy	Precision	Recall	F1 score	ROC_AUC score
MLP (DL)	0.75	0.31	0.06	0.10	0.51
CNN (DL)	0.76	0.55	0.03	0.05	0.51
Random forest classifier	0.77	0.56	0.07	0.13	0.53
Logistic regression	0.77	0.51	0.13	0.21	0.55
Gradient boosting	0.77	0.55	0.20	0.29	0.57
KNN	0.75	0.44	0.24	0.31	0.57

CNN, convolutional neural network; DL, deep learning; KNN, *K*-nearest neighbor; MLP, multilayer perceptron; ROC_AUC, area under the receiver operating characteristic curve.

Table 7: Performance metrics of different models for debit card classification

Model	Accuracy	Precision	Recall	F1 score	ROC_AUC score
MLP (DL)	0.65	0.55	0.42	0.48	0.60
CNN (DL)	0.68	0.60	0.54	0.57	0.65
Random forest classifier	0.69	0.63	0.50	0.55	0.68
Logistic regression	0.68	0.59	0.56	0.58	0.66
Gradient boosting	0.70	0.68	0.42	0.52	0.65
KNN	0.61	0.50	0.50	0.51	0.60

CNN, convolutional neural network; DL, deep learning; KNN, *K*-nearest neighbor; MLP, multilayer perceptron; ROC_AUC, area under the receiver operating characteristic curve.

known for its ability to handle complex datasets and mitigate overfitting, also showed strong performance in debit card classification, similar to its success in credit card classification tasks. Logistic regression, although a linear model, performed adequately but slightly less effectively than some other models in capturing the nuances of the data. Gradient boosting algorithms stood out for their strong performance, leveraging an ensemble of weak learners to improve predictive accuracy. Lastly, KNN demonstrated a balanced performance with moderate evaluation parameters, showcasing its reliability in classification tasks.

VII. Discussion

a. State prediction

The random forest algorithm emerges as the top-performing model for predicting the variable "State" in financial transactions, demonstrating exceptional accuracy and reliability. With an average accuracy of 99.99% across all states, the random forest model showcases its effectiveness in accurately determining the geographical origin of financial transactions. This superior performance can be attributed to numerous factors inherent to the random forest algorithm. First, random forest excels in handling complex datasets with high dimensionality, making it well-suited for modeling the intricate relationships between various features and the target variable. By aggregating predictions from multiple decision trees, random forest mitigates overfitting and generalizes well to unseen data, resulting in robust and reliable predictions. Additionally, the ensemble nature of random forest enables it to capture nonlinear relationships and interactions among features, enhancing its predictive power [37,38]. Techniques such as Bayesian optimization, as discussed by Surono et al. [39], could further refine the hyperparameters, enhancing model performance and efficiency.

b. Predicting state and year simultaneously

The random forest algorithm arose as the best-performing model among the evaluated regression algorithms for predicting both "Year" and "State" variables in financial transactions. Its superior performance, characterized by the lowest MSE and MAE values, highlights its effectiveness in making accurate predictions. The random forest algorithm's robustness lies in its ability to handle complex datasets with high dimensionality, capturing intricate relationships between variables. By leveraging an ensemble of decision trees, random forests can mitigate overfitting and enhance predictive accuracy, making it well-suited for modeling financial transaction data. However, despite its remarkable performance, the random forest algorithm encounters challenges in certain scenarios. One notable challenge is the interpretability of the model's predictions. Due to its ensemble nature and the complexity of decision trees, interpreting the underlying logic behind individual predictions can be challenging, limiting the model's transparency and understandability. Moreover, the scalability of random forest can be a concern when dealing with massive volumes of financial transaction data. While random forest generally performs well with moderate-sized datasets, scaling it to handle Big Data efficiently requires distributed computing frameworks and specialized infrastructure. Distributed computing frameworks, such as Apache Spark and Hadoop, facilitate the parallel processing of large datasets by distributing the data and computation across multiple nodes in a cluster. This parallelism significantly reduces the time required for

training and prediction by leveraging the combined processing power of multiple machines. Specialized infrastructure, including high-performance computing clusters and cloud-based services such as AWS EMR or Google Cloud Dataproc, provides the necessary resources to handle the increased memory and computational demands of large-scale random forest models. These frameworks and infrastructures also support fault tolerance and scalability, ensuring that the system can handle growing data volumes and recover from node failures without data loss [22]. Addressing these challenges necessitates a holistic approach involving advanced techniques in model interpretation, hyperparameter optimization, and scalability enhancements [40].

Strategies to optimize the performance of a RandomForestClassifier model, especially when dealing with imbalanced datasets or high-dimensional feature spaces, include the following:

- Handling class imbalance: Implementation of techniques such as synthetic minority oversampling technique (SMOTE) or adaptive synthetic sampling (ADASYN) is undertaken to balance the class distribution. Additionally, adjusting class weights or using cost-sensitive learning can help the model focus more on minority classes.
- Feature engineering: Carefully selecting and engineering features can significantly enhance model performance. Techniques such as principal component analysis (PCA) can reduce dimensionality while retaining important information, making it easier for the model to identify relevant patterns.
- Hyperparameter tuning: Optimizing hyperparameters such as the number of trees, maximum depth, and minimum sample split through techniques such as grid search or random search can lead to better model performance.
- Cross-validation: Using cross-validation ensures that the model's performance is consistent and not overly reliant on a specific subset of the data, leading to more robust and generalizable results.
- Ensemble methods: Combining random forest with other models through stacking or blending can improve predictive accuracy and robustness, especially for complex datasets.

c. Credit and debit card classification

The assessment of various models for credit and debit card classification provided valuable insights into their performance across different evaluation metrics. Among the algorithms evaluated, the random forest classifier emerged as a top-performing model for both credit and debit card classification tasks, exhibiting strong performance across multiple metrics.

For credit card classification, random forest demonstrated an accuracy of 0.77, a precision of 0.56, and an F1 score of 0.13, highlighting its effectiveness in accurately classifying credit card transactions. This exceptional performance is credited to its resilience and adaptability in handling complex datasets with high dimensionality, as well as its ability to capture nonlinear relationships among features. Similarly, for debit card classification, random forest exhibited an accuracy of 0.69, a precision of 0.63, and an F1 score of 0.55, showcasing its effectiveness in accurately classifying debit card transactions. Once again, random forest's robustness and versatility were instrumental in achieving strong performance, enabling it to mitigate overfitting and to generalize well to unseen data.

While random forest emerged as the top-performing model for both credit and debit card classification, other models, such as gradient boosting and logistic regression, also showed competitive results. Gradient boosting demonstrated high precision and a good F1 score, indicating its effectiveness in accurately classifying transactions. Gradient boosting works by building an ensemble of decision trees sequentially, whereby each tree attempts to correct the errors of the previous ones. This iterative approach enhances the model's ability to capture complex patterns and interactions within the data. Distinctive elements include the ability of gradient boosting to focus on difficult-to-classify examples in each iteration, resulting in improved prediction accuracy for credit and debit card transactions. Its learning rate and number of boosting stages were fine-tuned to optimize performance, ensuring that each subsequent tree was built upon the residual errors of the previous one, leading to a more precise model [38,39]. Logistic regression, although slightly less effective, still showcased utility in both classification tasks. Despite the promising performance of these models, challenges remain in achieving high recall scores, particularly in debit card classification, suggesting the need for further optimization and fine-tuning. Moving forward, future research endeavors should focus on enhancing model performance, particularly in terms of recall and F1 score, to better address the challenges associated with fraud detection in both credit and debit card transactions. Additionally, exploring ensemble methods and incorporating advanced feature engineering techniques may further enhance the predictive

accuracy and robustness of the models in real-world applications within the financial sector.

While DL techniques, including DNNs, were explored for predicting credit card and debit card transactions, they encountered challenges such as overfitting and interpretability issues. The inherent complexity of DL models limited their effectiveness in this specific task, highlighting the importance of choosing appropriate algorithms based on dataset characteristics and interpretability requirements.

The lower count in the credit card category compared to other categories can be attributed to several factors, as follows:

- Differences in transaction volume: Credit card transactions may be less frequent in certain datasets compared to debit card transactions. Debit cards are often used for everyday purchases, resulting in a higher volume of transactions and more data points for analysis.
- Reporting practices: Financial institutions might have different thresholds or practices for reporting suspicious credit card transactions compared to debit card transactions. This can lead to a lower number of reported suspicious credit card activities.
- Data collection bias: The dataset might reflect a bias in the types of transactions reported or collected. For example, certain types of suspicious activities might be more prevalent or more easily detected in debit card transactions, leading to a higher count in that category.

d. Application to strengthen AML in real world

Financial institutions can incorporate the advanced ML models developed in this study into their existing AML systems through several strategic steps. Initially, these institutions need to perform a thorough assessment of their current AML infrastructure to identify integration points where ML models can be most effective. The ML models can be deployed as an additional layer of defense within existing transaction-monitoring systems. By doing so, the models can analyze transaction data in real time, flagging potentially suspicious activities that may not be detected by traditional rule-based systems.

To achieve seamless integration, financial institutions can utilize application programming interfaces (APIs) to facilitate communication between the new ML models and their existing systems. APIs allow for the secure transfer of data and ensure that the models receive the necessary information for analysis. Furthermore, institutions can use middleware solutions to bridge any compatibility gaps between the new models and legacy systems, ensuring smooth data flow and operational consistency.

An essential aspect of this integration is the training and calibration of ML models using historical transaction data. Financial institutions should leverage their extensive databases to fine-tune the models, enhancing their accuracy and reducing false positives. This process involves continuous monitoring and adjustment of the models to adapt to emerging money-laundering tactics.

Additionally, institutions must establish protocols for handling alerts generated by the ML models. These protocols should define the procedures for investigating and escalating alerts, ensuring that suspicious activities are promptly and thoroughly examined. Integrating these models also necessitates regular collaboration between data scientists, AML compliance officers, and information technology (IT) personnel to maintain the system's effectiveness and address any technical challenges.

By incorporating ML models alongside existing AML systems, financial institutions can significantly enhance their ability to detect and prevent money laundering. This integration not only improves the efficiency and accuracy of transaction monitoring but also strengthens the overall security framework, helping institutions stay ahead of evolving financial crimes.

e. Challenges in implementing AI/ML in AML

One major challenge in integrating ML models into AML systems is the availability and quality of data. Financial transaction data may suffer from incompleteness, inaccuracy, or bias, which can adversely affect the performance of ML models. Additionally, accessing diverse and relevant data sources can be difficult due to regulatory constraints and concerns regarding data privacy. Another challenge is ensuring regulatory compliance with stringent requirements such as the BSA and the Anti-Money Laundering Act (AMLA). ML models must effectively detect suspicious activities while minimizing false positives and adhering to regulatory guidelines. Ensuring model transparency, fairness, and accountability is crucial for meeting regulatory compliance standards.

f. Limitations of relying on historical data

Historical data in financial crime analysis often lack timeliness, failing to reflect emerging patterns and

evolving tactics in money-laundering schemes. This deficiency is compounded by biases inherent in the data due to factors such as underreporting of suspicious activities, changes in reporting regulations, and variations in data collection practices over time, thereby leading to skewed predictions by ML models. Moreover, historical datasets frequently lack the contextual information necessary for a comprehensive understanding of reported suspicious activities, posing challenges in accurately interpreting patterns or trends. This is particularly critical given the dynamic nature of financial systems, which are influenced by factors such as economic conditions, technological advancements, and regulatory changes, all of which may not be adequately captured in historical data.

VIII. Future Work

a. Real-time monitoring system

Developing a real-time monitoring system for detecting suspicious transactions presents several challenges and requires a holistic approach, as follows:

- Data collection and preprocessing: The foundation of a real-time monitoring system lies in the aggregation of data from various sources. This includes transaction logs, customer profiles, regulatory watchlists, and external data such as market trends and geopolitical events. The integration of these diverse data sources into a centralized platform is crucial for comprehensive analysis. Once collected, the data undergo preprocessing, which includes cleaning to remove noise and inconsistencies, normalizing to standardize formats, and performing feature engineering. Feature engineering involves the extraction and transformation of relevant features from raw data to enhance the predictive power of the ML models. This step ensures the accuracy and consistency of the data used for model training, ultimately improving the reliability of the monitoring system.
- Real-time monitoring infrastructure: Building a scalable and robust infrastructure is essential for processing and analyzing data in real time. This infrastructure can leverage advanced technologies such as stream-processing frameworks, including Apache Kafka and Apache Flink, which are designed to handle high-throughput data streams with low latency. Additionally, cloud-based services such as AWS Kinesis and Google Cloud Pub/Sub provide scalable solutions for real-time data ingestion, storage, and processing. These technologies enable the continuous flow and immediate analysis of transaction data, which is critical for detecting suspicious activities as they occur.
- Rule-based filtering: In addition to ML models, rule-based filters play a vital role in a real-time monitoring system. These filters are designed to flag transactions that meet specific criteria for suspicious activity. The rules can be derived from regulatory requirements, industry standards, known fraud patterns, and institution-specific risk profiles. Implementing these filters helps in the preliminary identification of potentially suspicious transactions, providing an additional layer of scrutiny before ML analysis.
- Model deployment: Deploying ML models into a production environment for real-time monitoring is a critical step. These models, trained on historical and live transaction data, must be integrated seamlessly with existing transaction-processing systems. This integration allows the models to automatically analyze incoming data streams and identify anomalies or patterns indicative of money laundering. The deployment process involves setting up APIs for model inference, ensuring robust and low-latency communication between the models and the transaction systems.
- Alerting mechanism: An effective alerting mechanism is essential for timely notification of detected suspicious transactions. This system should be capable of generating alerts based on predefined severity levels and escalation procedures. The alerts should be configurable, allowing compliance officers and fraud analysts to prioritize and investigate high-risk transactions promptly. Additionally, the system should support various notification channels, such as emails, short message service (SMS), and dashboard alerts, to ensure immediate action.
- Feedback loops: To maintain and enhance the performance of the real-time monitoring system, establishing feedback loops is crucial. These loops involve continuous evaluation of model performance and incorporating feedback from fraud analysts and compliance officers. Regular model retraining using the latest transaction data helps in adapting to new fraud patterns and regulatory changes. Additionally, model calibration and performance monitoring ensure that the models remain accurate and effective over time. This iterative process of feedback and improvement is vital for keeping the monitoring system robust and responsive to emerging threats.

b. User-friendly interface for AML practitioners

Creating a user-friendly interface for AML practitioners involves using various technologies and strategies. Frontend development is crucial, utilizing modern frameworks such as React.js or Angular.js to design an intuitive and responsive user interface (UI). Incorporating industry-standard design principles and UI/user experience (UX) best practices ensures clear navigation and user-friendly interactions. Furthermore, dashboard development is essential, implemented using visualization libraries such as D3.js or Chart.js to display key metrics and alerts in a visually appealing and comprehensible manner. Real-time data-streaming technologies such as Apache Kafka or Amazon Kinesis facilitate timely updates, enhancing the dashboard's utility.

Role-based access control (RBAC) is another critical aspect, requiring the implementation of authentication and authorization mechanisms. Technologies such as JSON Web Tokens (JWT) enable secure user management, with frameworks such as Firebase authentication ensuring seamless integration. Alert management and case systems are developed using database technologies such as MongoDB or PostgreSQL, storing and managing alert data efficiently. Workflow automation tools such as Camunda or Apache Airflow streamline case management processes, reducing manual effort and enhancing productivity.

Effective search and filtering functionalities are integrated using technologies such as Elasticsearch or Apache Solr for fast and efficient search operations. Advanced filtering options are implemented using state management libraries such as Redux or MobX in frontend applications. Workflow automation is further enhanced through the integration of business process management (BPM) tools such as Activiti or Bonita BPM. Additionally, integration with robotic process automation (RPA) platforms such as UiPath or Automation Anywhere enables end-to-end automation of repetitive tasks, improving efficiency.

For compliance reporting, custom reporting tools are developed using business intelligence (BI) platforms such as Tableau or Power BI. These tools generate compliance reports, providing insights into efforts related to regulatory adherence and financial crime mitigation. Comprehensive documentation is provided using tools such as Swagger or GitBook, facilitating ease of use and troubleshooting. Multichannel support through ticketing systems such as Zendesk or Freshdesk, coupled with live chat solutions such as Intercom or LiveChat, ensures prompt assistance and user satisfaction. Moreover, feedback mechanisms using survey tools such as SurveyMonkey or Typeform, along with analytics platforms such as Google Analytics or Mixpanel, enable continuous improvement based on user input and usage analytics. Through the integration of these technologies and strategies, AML practitioners can access a user-friendly interface that enhances their efficiency and effectiveness in combating financial crimes.

Additionally the following tools and techniques can be considered in future research.

1. Explainable AI (XAI) techniques: The manuscript does not incorporate XAI techniques in the models to promote transparency and build trust with regulators and stakeholders. Future work should focus on integrating XAI methods to ensure that AML models are interpretable and their decisions can be understood by human analysts.
2. Federated learning systems: The proposed method fails to create a federated learning system that allows multiple financial institutions to collaborate and share data to improve their AML models while protecting customer privacy. Developing such a system could significantly enhance the collective ability to detect and prevent money-laundering activities while maintaining data security.
3. AI/ML-powered risk assessment tool: To identify high-risk individuals or entities, it is crucial to develop an AI/ML-powered risk assessment tool to analyze customer data. Such a tool would help in proactively managing and mitigating risks associated with potential money-laundering activities.

   Financial institutions can customize risk criteria based on their specific regulatory requirements and risk tolerance levels. The AI/ML-powered tool allows the creation of tailored risk profiles, incorporating various factors such as transaction volume, frequency, origin and destination of funds, and historical risk scores. This customization ensures that the tool aligns with the institution's risk management strategy and regulatory obligations.

   The AI/ML-powered risk assessment tool can be seamlessly integrated with existing financial systems, such as customer relationship management (CRM) systems, transaction processing platforms, and regulatory compliance software. This integration ensures a smooth flow of data and enhances the overall efficiency of risk management processes. Financial institutions can leverage their existing infrastructure while benefiting from advanced AI/ML capabilities.

4. Analyzing unstructured data using natural language processing (NLP): While recognizing the potential benefits of incorporating NLP methods into AML efforts, such as extracting insights from textual sources such as financial documents, emails, and chat logs, this study primarily focuses on structured financial transaction data. Given the complexity and breadth of the AML domain, including both structured and unstructured data sources, NLP techniques could indeed provide valuable insights into detecting and mitigating financial crimes. However, due to the scope limitations of this paper, which primarily examines structured data analysis using ML algorithms, an in-depth exploration of NLP techniques falls outside our current research. Future studies could benefit from integrating NLP methods to analyze unstructured data sources in conjunction with structured data analysis, offering a more comprehensive approach to AML detection and prevention.

IX. Conclusion

The integration of AI/ML algorithms presents a promising avenue for fortifying AML frameworks, as evidenced by the findings and analysis presented in this research paper. The research delved into the intricate realm of financial transactions, using cutting-edge ML techniques to predict states and years, in addition to classifying credit and debit card activities. Central to the findings was the random forest classifier's remarkable accuracy, achieving an unprecedented 99.99% accuracy in state prediction. This robust performance, as depicted in the confusion matrix, not only underscores the model's reliability but also its crucial role in distinguishing between fraudulent and nonfraudulent transactions—a vital aspect in financial analytics and fraud detection. Simultaneous prediction of year and state added another layer of complexity, showcasing gradient boosting's competitive edge in explaining data variance with the highest R^2 value among the algorithms. Meanwhile, the random forest algorithm's low MSE highlighted its superior predictive accuracy in handling such multifaceted tasks. These results mark significant theoretical contributions, emphasizing the efficacy of ensemble methods such as random forest and gradient boosting in modeling intricate patterns within financial data. The study's practical implications are profound, particularly in enhancing fraud detection mechanisms for financial institutions. By integrating these advanced ML models into existing systems, institutions can bolster their ability to accurately predict state origins and transaction years, as well as classify credit and debit card activities. However, the research is not devoid of limitations. The study does not delve into potential biases in the data or how these biases might affect the performance of the algorithms, which is crucial in the context of fair and unbiased fraud detection systems. Challenges such as data availability and computational costs are acknowledged, presenting avenues for future exploration and refinement. Looking ahead, the study calls for actionable, practical scholarship that leverages these advanced ML techniques to address real-world challenges in fraud prevention and financial analytics. The demonstrated success of ensemble methods and DL models serves as a catalyst for further advancements and innovations in the field. Future work based on the findings of this paper could focus on several key areas to further enhance the application and impact of ML in financial analytics and fraud detection. Incorporating XAI techniques such as SHAP or LIME can improve model interpretability, fostering trust and understanding among stakeholders. Integrating additional data sources beyond transactional data, such as customer behavior and socioeconomic factors, can provide a more comprehensive context for analysis, leading to richer insights and improved predictive accuracy

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

This research was supported by the Centre for Advanced Modelling and Geospatial Information Systems (CAMGIS), Faculty of Engineering and Information Technology, the University of Technology Sydney, NSW, Australia; and in part by the Researchers Supporting Project number RSP2024 R14, King Saud University, Riyadh, Saudi Arabia.

References

[1] Kute D V, Biswajeet P, Nagesh S, & Abdullah A, "Deep learning and explainable artificial intelligence techniques applied for detecting money laundering–a

critical review", *IEEE access* 9: 82300-82317, June 2021.

[2] Abid D, Rahmatullah W A, Alif M A F, & Rinci K H, "Penerapan Metode K-Means Clustering Untuk Analisa Penjualan Komoditas Toko Tani Indonesia", *KERNEL: Jurnal Riset Inovasi Bidang Informatika dan Pendidikan Informatika* 3, no. 2: 25-30, Oct 2022.

[3] Kavisha M S, "Anti Money Laundering: Proactive involvement and perception of Internal Auditors in Anti-Money Laundering Compliance Review", PhD diss., GUJARAT TECHNOLOGICAL UNIVERSITY AHMEDABAD, Feb 2024.

[4] Omri R, "Applying supervised machine learning algorithms for fraud detection in anti-money laundering", *Journal of Modern Issues in Business Research* 1, no. 1: 14-26, Dec 2021.

[5] Zhiyuan C, Dinh V K L, Ee N T, Amril N, Ettikan K K, & Kim S L, "Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: a review", *Knowledge and Information Systems* 57: 245-285, Feb 2018.

[6] Ítalo D G, Luiz H A C, & Erick G M, "Graph Neural Networks Applied to Money Laundering Detection in Intelligent Information Systems", In *Proceedings of the XIX Brazilian Symposium on Information Systems*, pp. 252-259, May 2023.

[7] Mark W, Giacomo D, Jie C, Daniel K I W, Claudio B, Tom R, & Charles E L, "Anti-money laundering in bitcoin: Experimenting with graph convolutional networks for financial forensics", *arXiv preprint arXiv:1908.02591,* July 2019

[8] Rasmus I T J, & Alexandros I, "Fighting money laundering with statistics and machine learning", *IEEE Access* 11: 8889-8903, Jan 2023.

[9] Charitou C, Simo D, & Artur D G, "Synthetic data generation for fraud detection using gans", *arXiv preprint arXiv:2109.12546*, Sept 2021.

[10] Fredrik J, & Martin J, "Finding Money Launderers Using Heterogeneous Graph Neural Networks", *arXiv preprint arXiv:2307.13499*, July 2023.

[11] Xiong K, Binhui P, Yang J, & Tiying L, "A hybrid deep learning model for online fraud detection", In *2021 IEEE International Conference on Consumer Electronics and Computer Engineering (ICCECE)*, pp. 431-434, Jan 2021.

[12] Jingguang H, Yuyun H, Sha L, & Kieran T, "Artificial intelligence for anti-money laundering: a review and extension", *Digital Finance* 2, no. 3: 211-239, June 2020.

[13] Ashwini K, Sanjoy D, Vishu T, Rabindra N S, & Ankush G, "Analysis of classifier algorithms to detect anti-money laundering", *Computationally intelligent systems and their applications*: 143-152, Apr 2021.

[14] Ahmed N B, Almohammady A, Mohamed S F, & Kamal R R, "Combating Financial Crimes with Unsupervised Learning Techniques: Clustering and Dimensionality Reduction for Anti-Money Laundering", *arXiv preprint arXiv:2403.00777*, Apr 2024.

[15] Alkhalili M, Mahmoud H Q, & Fadi A, "Investigation of applying machine learning for watch-list filtering in anti-money laundering", *iEEE Access* 9: 18481-18496, Jan 2021.

[16] William G, & Athenia B S, "Anti-money laundering and customer due diligence: empirical evidence from South Africa", *Journal of Money Laundering Control* 26, no. 7: 224-238, Dec 2023.

[17] Charanjit S, & Wangwei L, "Can artificial intelligence, RegTech and CharityTech provide effective solutions for anti-money laundering and counter-terror financing initiatives in charitable fundraising", *Journal of Money Laundering Control* 24, no. 3: 464-482, July 2021.

[18] Boris K, Evgenii V, Alexander D, & Antoni W, "Interpretable Machine Learning forFinancial Applications", In *Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook*, pp. 721-749. Cham: Springer International Publishing, Aug 2023.

[19] Lucas S G, André L K, Platão G T N, Davenilcio L S, & Taciana M, "Anti-money laundering and financial fraud detection: A systematic literature review", *Intelligent Systems in Accounting, Finance and Management* 29, no. 2: 71-85, May 2022.

[20] Wei-Yu C, Shing-Han L, & Yung-Hsin W, "Research on Natural Language Processing in Financial Risk Detection", In *Cognitive Cities: Second International Conference, IC3 2019, Kyoto, Japan, September 3–6, 2019, Revised Selected Papers 2*, pp. 448-455. Springer Singapore, June 2020.

[21] Abdul K L, & Leyla, "Anomaly Detection in Financial Transaction Time Series Data", June 2023.

[22] Farman A, & Pradeep S, "Big Data Analytics in Financial Econometrics", *Current Studies in Social Sciences 139*, Dec 2022.

[23] Nadia P, Mirko Z, Fabio M, Muhammad Z S, & Stefano F, "Detecting anomalous cryptocurrency transactions: an aml/cft application of machine learning-based forensics", *arXiv preprint arXiv:2206.04803*, June 2022.

[24] Mark L, "Predicting money laundering using machine learning and artificial neural networks algorithms in banks", *Journal of Applied Security Research* 19, no. 1, Jan 2024.

[25] Martin J, Anders L, Ragnar B H, Geir Å, & Johannes L, "Detecting money laundering transactions

with machine learning", Journal of Money Laundering Control 23, no. 1 173-186, Jan 2020.

[26] Wai W L, Mohanad S, Siamak L, & Marius P, "Inspection-L: A Self-Supervised GNN-Based Money Laundering Detection System for Bitcoin", Mar 2022.

[27] Sizheng W, & Suan L, "Financial Anti-Fraud Based on Dual-Channel Graph Attention Network", Journal of Theoretical and Applied Electronic Commerce Research 19, no. 1: 297-314, Feb 2024.

[28] Alotibi J, Badriah A, Tahani A, Hosam A, & Abdullah B, "Money Laundering Detection using Machine Learning and Deep Learning", International Journal of Advanced Computer Science and Applications 13, no. 10, Jan 2022.

[29] Zhenfeng S, Muhammad N A, & Akib J, "Comparison of Random Forest and XGBoost Classifiers Using Integrated Optical and SAR Features for Mapping Urban Impervious Surface", Remote Sensing 16, no. 4: 665, Feb 2024.

[30] Nevine L, Mohammed A R, & Amr E M S, "Survey of machine learning approaches of anti-money laundering techniques to counter terrorism finance", In Internet of Things—Applications and Future: Proceedings of ITAF 2019, pp. 73-87. Singapore: Springer Singapore, Apr 2020.

[31] Haobo Z, Junyuan H, Fan D, Steve D, Liangjie X, & Jiayu Z, "A privacy-preserving hybrid federated learning framework for financial crime detection", arXiv preprint arXiv:2302.03654, Feb 2023.

[32] Li Y, & Abdallah S, "On hyperparameter optimization of machine learning algorithms: Theory and practice", Neurocomputing 415 295-316, Nov 2020.

[33] Guy S H, Hong K K, Ronil V C, Amir H R, Shiwei H, Mark B, Michael J L, & Hamed A, "Peering into the black box of artificial intelligence: evaluation metrics of machine learning methods", American Journal of Roentgenology 212, no. 1 38-43, Jan 2019.

[34] Abhishek V T, "Comparative assessment of regression models based on model evaluation metrics", International Research Journal of Engineering and Technology (IRJET) 8, no. 09 2395-0056, Sep 2021.

[35] Željko Đ V, "Classification model evaluation metrics", International Journal of Advanced Computer Science and Applications 12, no. 6 599-606, July 2021.

[36] Alaa T, "Classification assessment methods", Applied computing and informatics 17, no. 1 168-192, July 2020.

[37] Xiao W, Meiqi Z, Deyu B, Peng C, Chuan S, & Jian P, "Am-gcn: Adaptive multi-channel graph convolutional networks", In Proceedings of the 26th ACM SIGKDD International conference on knowledge discovery & data mining, pp. 1243-1253, Aug 2020.

[38] Sugiyarto S, M. Yahya F A, Anggi S, Dyiyah K E A, & Aris T, "Comparison of CNN Classification Model using Machine Learning with Bayesian Optimizer", HighTech and Innovation Journal, 4(3), 531-542, Sept 2023.

[39] Surono S, Yahya M F A, Anggi S, Dyiyah K E A, & Aris T, "Comparison of CNN Classification Model using Machine Learning with Bayesian Optimizer", HighTech and Innovation Journal, 4(3), 531-542, Sept 2023.

[40] Dibs H, Abu Dabous S, Shaaban M, & Marzouk M, "Multi-Fusion Algorithms for Detecting Land Surface Pattern Changes Using Multi-High Spatial Resolution Images and Remote Sensing Analysis", Remote Sensing, 13(11), 2098, June 2023.