NOVA IMS Information Management School

MDSAA

Master Degree Program in

Data Science and Advanced Analytics

Money Laundering and Fraud detection under concept drift

A Systematic Review

Simão Ortigão Baptista

Master Thesis

presented as partial requirement for obtaining a Master's Degree in Data Science and Advanced Analytics

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

Money Laundering and Fraud detection under concept drift

A Systematic Review

by

Simão Ortigão Baptista

Master Thesis presented as partial requirement for obtaining the Master's degree in Data Science and Advanced Analytics, with a specialization in Business Analytics

Supervised by

Bruno Damásio, PhD, NOVA Information Management School

July, 2025

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledged the Rules of Conduct and Code of Honor from the NOVA Information Management School.

[Lisbon, 10th July 2025]

ACKNOWLEDGEMENTS

As a new academic milestone approaches, I can't express my gratitude for all the support I've received throughout this journey.

To Professor Bruno Damásio, PhD, who guided me throughout this thesis, sharing constructive ideas and much-requested feedback that were essential to the development of this study.

To my family, far away but always present, who from an early age instilled in me the principles and values that characterize me and for whom all the effort was worthwhile. None of this would have made sense without their support. They were always in my thoughts and were my motivation.

To my friends, whom I could always count on to vent and relax. It was moments like these that also gave me the clarity I needed to accomplish all of this.

To my coworkers, who welcomed me from the beginning and with whom I exchanged so many insights with a direct and indirect impact on this study.

To everyone, my sincere gratitude.

ABSTRACT

Financial Institutions are pivotal yet vulnerable agents in societies. Two of the threats they face are money laundering and fraud, requiring thorough risk mitigation strategies and quick response to avoid financial, operational and reputational costs. The evolving dynamics of these threats make it harder for traditional, rule-based detection systems to follow up, becoming obsolete over time. In contrast, machine learning's ability to handle complex, non-linear patterns and learn temporal sequences makes it a promising alternative. This study presents a systematic literature review that synthesizes a collection of peer-reviewed original works published between 2014 and 2024, identifying and categorizing machine learning approaches to each domain, with an emphasis on understanding how concept drift is handled. The analysis highlights prevailing algorithmic strategies, comparative insights between the two domains, knowledge gaps and future research directions to address them. The findings contribute to a structured understanding of machine learning applications in financial crime detection and strengthen decision-making towards more adaptive, data-driven detection systems.

KEYWORDS

Money Laundering; Fraud; Financial Institution; Machine Learning; Concept Drift

Sustainable Development Goals (SDG):

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ANFIS Adaptive Neuro-Fuzzy Inference System

ANN Artificial Neural Networks

CDD Customer Due Diligence

CNN Convolutional Neural Networks

DT Decision Tree

FATF Financial Action Task Force

FI Financial Institution

GAT Graph Attention Network

GIN Graph Isomorphism Network

GRU Gated Recurrent Units

KNN K-Nearest Neighbors

KYC Know Your Customer

LPA Label Propagation Algorithm

LR Logistic Regression

LSTM Long Short-Term Memory

ML Machine Learning

MLP Multilayer Perceptron

NB Naïve Bayes

RF Random Forests

RNN Recurrent Neural Networks

SAR Suspicious Activity Reports

SLR Systematic Literature Review

SVM Support Vector Machines

TM Transaction Monitoring

1. INTRODUCTION

Financial Institutions (FI) play a pivotal role by channeling financial resources from surplus agents to those in deficit, enabling a more efficient reallocation of capital that promotes economic growth and stability. Technological advancements boosted their centrality to economies, allowing more transactions and, with it, more transactional data volume and complexity, making it harder for fraud detection systems to navigate through it. That poses critical challenges for FIs, who became highly vulnerable to illicit actions from criminals that exploit traditionally rule-based detection mechanisms by designing new ways to go around them faster than the systems themselves can adapt to, largely affecting FIs' operations, finances, reputation and trust from their clients. The static nature of these mechanisms is particularly problematic given the concept drift phenomenon, where the statistical properties of financial crime data change over time and require appropriate mitigation strategies. Given these escalating challenges and the limitations of traditional approaches, machine learning (ML) emerges as a suitable alternative, providing scalable and adaptable solutions by learning complex, non-linear patterns and relationships in transactional datasets that would have gone unnoticed otherwise. Moving towards these emerging approaches becomes a need as regulatory and compliance pressures intensify, demanding better and faster responses to fraudulent activity.

The growing attention towards the application of ML to these topics has led to several works over the years. Previous Systematic Literature Reviews (SLR) have examined the use of ML in financial crime detection but either concentrate exclusively on fraud or money laundering, treating them as independent problems despite their potential overlap. Other SLRs that attempt to bridge the two topics, even offering a conceptual framework for joint detection efforts, provide limited attention to ML application (Goecks et al., 2020). Moreover, previous SLRs tend to focus on challenges like class imbalance and label availability, not centering their study on concept drift. Considering previous progress and the identified knowledge gap, this study will develop an SLR that will try to answer the following research question: **How is machine learning being applied to detect money laundering compared to fraud in financial institutions under concept drift?**

To answer it, a few research objectives will be outlined. Firstly, the main purpose of the SLR is to review and synthesize the latest existing literature regarding the use of ML techniques in detecting both fraud and money laundering in FIs. Following that, the SLR aims to examine how concept drift is addressed within such approaches, identifying methods to adapt to changing data's statistical properties over time. Finally, to foster further investigation, the SLR aspires to point out relevant knowledge gaps and future work directions. Ultimately, the SLR will contribute to a clearer understanding of how ML research addresses money laundering and fraud under dynamic data conditions.

The structure of the SLR is as follows: **Section 2** provides background context to the reader across three key dimensions to understand the remaining of the work: FI's landscape, financial crime and ML principles; **Section 3** outlines the methodology adopted in the development of the SLR, namely the guiding framework, search strategy and inclusion criteria; **Section 4** presents the results of the SLR along a quantitative analysis of publication patterns regarding the two topics and a qualitative synthesis of the works included; **Section 5** offers a critical discussion of the findings of the SLR, linking them to the proposed research question, their practical implications to relevant stakeholders, limitations of the study and future work directions; **Section 6** concludes the SLR by summarizing the main insights and their relevance.

2. BACKGROUND CONTEXTUALIZATION

2.1. FINANCIAL INSTITUTIONS LANDSCAPE

Economies are driven by the interaction between supply and demand, thriving when these two forces meet efficiently (Madison, 2023). At the heart of the resulting transactions lie FIs, and more precisely Monetary FIs. These agents, which include commercial banks and similar organizations, serve as intermediaries of such operations by performing essential economic functions, namely accepting deposits from savers – who prefer to set funds aside for future use – and providing loans to borrowers – who would rather increase their consumption today in exchange for future repayment with interest (European Central Bank, 1998). That interest represents the "price" of money and serves not only as the primary source of revenue for Monetary FIs but also as a policy instrument to encourage or discourage borrowing and consumption within an economy. The crucial role of Monetary FIs in fostering economic development is not up to debate: by channeling financial resources from surplus to deficit agents, they contribute to a more efficient allocation of capital, promote long-term growth and economic stability by ensuring liquidity and absorbing market shocks. Thus, they end up being one of the backbones of both micro and macroeconomics' resilience and sustained growth (European Central Bank, 2012).

However, technological advancements and global growing interconnectivity pose a variety of challenges: the ease with which transactions are currently being held is leading to a huge increase in transactional data volume that needs to be handled carefully by these Monetary FIs, who need to keep up with that pace to ensure not only operational robustness but also their clients' protection from threats that may go unnoticed within the complex nature of the data (Financial Action Task Force, 2021).

So, despite their pivotal role, Monetary FIs are not immune to exploitation from bad-intentioned people: their scale, complexity and centrality to economies make them attractive targets for illicit actors, whose actions can damage institutional integrity and affect innocent individuals, markets and entire financial systems (Levi, 2002). Among the most pervasive threats faced by these Monetary FIs are money laundering and fraud, that exploit their gaps and vulnerabilities. Fraud usually involves deception for personal or financial gain and can take on many forms, from which credit card fraud gains relevance to the agents in this study (Beals et al., 2015). On the other hand, money laundering is the process of disguising the true origins of illicitly obtained funds, often involving heterogeneous and complex layers of financial transactions to mislead authorities (Levi, 2002). While distinct in their dynamics and motivations, these illicit activities often intersect, as proceeds from fraudulent schemes tend to be laundered to appear legitimate in the eyes of regulators (Betron et al., 2012; Levi, 2002).

To address these threats, Monetary FIs are subject to stringent Compliance programs, designed and shaped by national legislation and guided by international frameworks like the

Financial Action Task Force (FATF) that provide recommendations to these agents. Among the processes of a Compliance program are (Financial Action Task Force, 2012):

- **Know Your Customer (KYC)**: the initial process of verifying a customer's identity before engaging in any serious financial relationship by collecting official identification documents, assuring personal details and understanding the nature and purpose of the relationship;
- **Customer Due Diligence (CDD)**: the procedure of a customer's risk assessment based on factors like the source of the funds, transaction behavior and geographic exposure to identify customers that urge enhanced supervision;
- **Transaction Monitoring (TM)**: the ongoing review of a customer's transactional behavior, analyzing activity across accounts to identify potentially suspicious patterns that may indicate fraudulent conduct;
- **Suspicious Transaction Reports (STR)**: a formal notification to be filed with a regulatory authority when activity that may involve funds derived from dubious origin or transactions designed to evade reporting requirements is detected.

Yet, these important processes often rely on rigid rule-based mechanisms that translate into large volumes of false positives and struggle to adapt to evolving criminal behavior. As financial crime becomes more sophisticated and complex, the limitations of traditional Compliance programs have become evident, which has been fueling the interest towards data-driven approaches, namely ML which offers the potential to enhance detection capabilities and respond more effectively to emerging threats (Financial Action Task Force, 2021).

2.2. UNDERSTANDING FINANCIAL CRIME: FRAUD AND MONEY LAUNDERING

Fraud and money laundering are two overlapping topics, many times related, that ultimately benefit some illicit agents through a set of complex and misleading financial operations to confuse legal and regulatory authorities (Betron et al., 2012; Levi, 2002).

According to United Nations 1988 Convention Against Illicit Traffic in Narcotic Drugs and Psychotropic Substances' Article 3, money laundering can be defined as "the conversion or transfer of property, knowing that such property is derived from any offence or offences […] for the purpose of concealing or disguising the illicit origin of the property or of assisting any person who is involved in the commission of such an offence or offences to evade the legal consequences of his actions" (United Nations, 1988). From a technical perspective, it is commonly described as a three-stage process: "first, the funds are introduced into the legitimate financial system (**placement**); next, complex financial transactions are carried out to camouflage the illegal source of the cash (**layering**); and lastly, legitimate assets can be

purchased using the illicit funds (**integration**)" (United Nations, 2021). Essentially, the goal is to disguise the criminal nature of such capital and make it look legitimate through a meticulous procedure of financial operations to deceive authorities. It is important for FIs to rapidly address laundering schemes as they cause harm at different dimension: empower criminal groups by allowing self-financing, undermines FIs trough corruption and compromises both political and social order (Levi, 2002). Laundering schemes come in several forms and shapes, for which such flexibility is required in detection systems to be ready to read between the lines and alert for suspicious activity independently of the form. Common techniques employed range from Smurfing, in which large amounts of cash are introduced into the financial system through several small deposits to avoid scrutiny, to Cash-intensive businesses, where the illicit funds are blended with legitimate revenues like those from restaurants or convenience stores for example, and Money Mules, accounts that serve as intermediaries to launder ill-gotten funds not always with full awareness of the owner (Levi, 2002; Rani et al., 2024). So, several dimensions come into play when dealing with such matters, such as transactional information (e.g. amount, direction of the operation, timestamp, etc.) but also customer attributes (e.g. type of customer, purpose of the relationship, origin of funds, etc.). Addressing money laundering is of great concern for global economies to preserve order and promote sustained economic growth but there is still a long way to go as it is estimated that 2%-5% of global GDP is laundered every year (United Nations, 2014). FIs play a pivotal role in this fight as they are one of the key entry points of illicit funds into the financial system and need to protect their interests and those of their customers. In response to that, governments and international bodies have enhanced regulatory pressure, prompting FIs to implement strict Compliance programs, tailored by FATF Recommendations' frameworks to conduct Customer Due Diligence, monitor suspicious activity and report anomalies through formal channels, shifting FIs' position from passive agents to active participants in the fight against organized financial crime (Betron, 2012).

On the other hand, fraud can be defined as "intentionally and knowingly deceiving the victim by misrepresenting, concealing, or omitting facts about promised goods, services, or other benefits and consequences that are nonexistent, unnecessary, never intended to be provided, or deliberately distorted for the purpose of monetary gain" (Beals et al., 2015). In the context of FIs, fraud can be originated from several sources from which a relevant distinction has to be made: Internal fraud, carried out by employees, and External fraud, attributed to customers or third parties (Sanusi et al., 2015). To this research, focus will be towards the latter. Frequent examples of External fraud include credit card fraud, relevant at both traditional and digital transactional environments in which the customer's credit card is used to make purchases or withdrawn funds typically without his consent, and identity theft, where there is a deliberate use of someone else's personal information to commit fraud – open accounts, access to credit, etc. (Sudjianto et al., 2012; Sanusi et al., 2015). This kind of fraud poses a threat to FIs, not only through direct monetary losses but also by compromising operational integrity and stakeholders' trust, eventually affecting shareholders, depositors, and borrowers and, consequently, the economy as a whole due to these institutions' crucial

role. Detecting fraudulent behavior is essential to FIs but fraudsters' creativity and adaptability to undercover such practices with new techniques makes that task harder, emphasized by the growing digital presence of transactional operations that allows for new technological ways of circumventing controls and at a much faster pace (Sanusi et al., 2015). To address the urgency of the topic, FIs need strong governance structures, with independent and competent boards, to reduce the likelihood of such fraudulent events (Wang et al., 2013). In fact, FIs are under increasing regulatory pressure to meet expectations, viewing fraud risk not just as an internal audit concern but rather as a core compliance issue, requiring formal programs and enhanced monitoring (Betron, 2012). However, traditional rule-based approaches to detect fraud may not capture the whole complex fraudulent network, falling behind emerging trends, emphasizing the limitations of such static detection systems (Sudjianto et al., 2010), urging for more sophisticated ways to address the issue.

The relationship between money laundering and fraud is well-supported in the context of FIs: laundering frameworks are inherently dependent on an underlying offense, such as fraud, to illicitly obtain the funds to be laundered (Levi, 2002). Fraud is not necessarily always the predicate offense that drives criminals to launder criminal proceedings but realizing their relationship is essential to understand the need to not analyze the topics independently but rather pursue forward-looking frameworks that address their interdependence through integrated detection and compliance strategies (Betron, 2012).

There are several characteristics in financial crime data that need to be dealt with carefully to accurately address the issue. Fraudulent activity involves navigating through vast volumes of high-dimensional transactional data, both in depth and width; fraudulent events tend to be rare, meaning a high class imbalance between fraudulent and non-fraudulent activity exists; it is also difficult to find real labeled datasets regarding financial transactions for confidentiality reasons, as the content is too sensitive both for the customers and the FIs (Lu et al., 2019; Sudjianto et al., 2012). Yet one of the most challenging factors is concept drift, the tendency for the statistical properties of financial data to change over time, as new fraudulent schemes and operations emerge. In essence, it occurs whenever the joint probability distribution P(X,y), where X represents the input features and y the target variable, fraudulent or non-fraudulent, shifts away, meaning that what once was true may not hold now (Lu et al., 2019). Statistically, it is typically addressed as structural breaks in time series models (Damásio et al., 2024). Intuitively, it makes any kind of static detection system obsolete over time, leaving FIs unprotected and vulnerable to criminals (Lu et al., 2019). In such a volatile and adversarial environment like that of financial crime, ML capabilities are crucial, offering more flexibility, scalability and data-driven insights to support decision making in the fight against criminals. In fact, such data-driven methods have been increasingly applied to detect fraudulent activity in other domains that share similar challenges – class imbalance and concept drift – such as public procurement (Lyra et al., 2022).

2.3. MACHINE LEARNING FOUNDATIONS

ML is a branch of Artificial Intelligence whose scope is to enable computers to learn from data and provide enhanced performance on a given task based on past experiences without being explicitly programmed to it (Pallathadkaa et al., 2023). By identifying patterns within datasets, ML algorithms can provide predictions, classifications or even detect anomalies, capabilities that could fall short if relying solely on manual analysis or traditional rule-based approaches, as these would not only not identify non-linear and complex relationships (Goodel et al., 2021) but also not have enough capacity to respond to the large transactional volume of the data (Pallathadkaa et al., 2023). ML's role is to augment the capacity of delivery in its various forms: from handling complexity to address large amounts of data, providing thoughtful insights to assist humans' decision-making (Goodel et al., 2021).

ML methods are commonly divided into three learning paradigms: Supervised, Unsupervised and Reinforcement. **Supervised Learning** is the process of learning from labeled instances of previous experience to produce generalized hypotheses that can predict unseen examples. Each instance of the dataset includes a known output label, from which a model of the distribution of class labels in terms of predictor features is expected to be built. The final classifier, based on such model, is then used to assign class labels to incoming instances, where only the predictor features are known. To ensure generalizability, the classifier should employ validation techniques to avoid overfitting, meaning the classifier would be excellent in predicting what it had already seen but a poor predictor of unseen examples (Kotsiantis, 2007). **Unsupervised Learning**'s purpose is to group collections of unlabeled instances together based on similarity measures. These methods operate on top of instances that don't have any predefined label attributed in the first place, relying solely on the harmony generated across the characterizing feature set to find natural groupings. As such, it is considered a form of exploratory data analysis, involving tasks like feature selection and proximity measurement that will lead up to the groupings of instances, also called clusters. These clusters are built on top of the assumption that instances within the same cluster are more similar among themselves than to instances from other clusters, where similarity usually is some distance metric (Jain et al., 2000). **Reinforcement Learning** is the problem faced by an agent that must learn behavior through experimental interactions with a dynamic environment, receiving scalar signals that guide its actions. At each step of an interaction, the agent understands the state of the environment, chooses an action and receives a reward along with a new state. The goal of the agent is to maximize the long-run cumulative reward function, having to balance exploration of unknown actions with exploitation of known rewarding ones by interacting with the environment. These methods offer greater adaptability compared to other traditional methods for its capacity to operate under conditions of uncertainty and partial observability (Kaelbling et al., 1996).

These three learning paradigms encapsulate a wide range of algorithmic strategies that vary in complexity, interpretability and adaptability. Given the cruciality of selecting the

appropriate algorithms and techniques to tackle both money laundering and fraud, and for the sake of structure of this research, the discussed methods in the literature have been organized into seven functional groups that reflect computational principles and detection capabilities: **Neural Networks and Deep Learning**; **Graph-based**; **Tree-based**; **Statistical and Probabilistic**; **Support Vector Machines**; **Ensemble methods**; **Clustering and Instance-based**. It is important to note that these functional groups are not necessarily mutually exclusive: some algorithms may exhibit characteristics that align with more than one group, but for analytical consistency each algorithm is assigned to the group that best reflects its dominant structure or use case within the financial crime detection landscape. The following paragraphs will explore each in more detail.

Neural Networks and Deep Learning techniques mark a foundational advancement in the ML field, enhancing the modelling of complex and non-linear patterns within data. These algorithms' roots lie in the essence of a **perceptron**, a computational unit inspired in the behavior of a human neuron by performing a weighted sum of inputs to be fed into a non-linear activation function so that a final output can be retrieved under a triggering signal (Pal et al., 1992). Building on this base, the **Multilayer Perceptron (MLP)** is a compilation of multiple layers of perceptrons, where the output of each perceptron of a layer is to serve as input to each perceptron of the subsequent layer, allowing the network to approximate non-linear functions and solve complex classification problems via backpropagation signals that iteratively adjust the connection weights (Pal et al., 1992). MLPs are the core structure of **Artificial Neural Networks (ANN)**, an algorithm widely resorted in the ML field for its pattern recognition, prediction and classification capabilities across a variety of domains, namely in financial crime detection (Abiodun et al., 2018). A specialized class of Deep ANNs are **Convolutional Neural Networks (CNN)**, meant to exploit spatial hierarchies within data through convolutional layers that learn local spatial features and pooling operations that reduce feature maps to control for model complexity, ideal for pattern recognition across grid-like structure datasets, such as images (Gu et al., 2018). Another class of Deep ANNs are **Recurrent Neural Networks (RNN)**, offering temporal modelling capabilities enabled by their internal feedback loops that allow the network to incorporate information from previous inputs into current computations, making it ideal for sequential data (Yu, 2024), such as the transactional data of FIs. To address the lack of long-term view limitation of traditional RNN, **Long Short-Term Memory (LSTM)** networks introduce memory cells and gate mechanisms to maintain and capture long-term dependencies across the data, augmenting the performance of its predecessor in long-term environments (Yu, 2024). A computationally cheaper version of LSTMs is **Gated Recurrent Units (GRU)** that blends the forget and input gates, ultimately reducing complexity while maintaining performance quality for sequence learning tasks (Yu, 2024). Lastly, the **Adaptive Neuro-Fuzzy Inference System (ANFIS)** combines neural networks with fuzzy logic, introducing fuzzy if-then rules through an adaptive network structure to model non-linear functions, designed to perform under uncertain environments (Jang, 1993).

Graph-based algorithms exploit the relational structure of graphs to learn from data where nodes – entities – and their mutual interactions – edges – are evidenced. A foundational unsupervised approach is the **Label Propagation Algorithm (LPA)**, which infers class membership through an iterative process based on the edges among the nodes of a graph dataset. An initial unique label is attributed to each node in the graph and at each iteration it assumes the label that is most common among its neighbors, converging to a consensus over time, standing out for its clarity, scalability and non-parametric nature (Garza et al., 2019). On the inductive representation side, **GraphSAGE** is an algorithm that learns how to generate vector representations for nodes based on the characteristics of their neighborhood, instead of considering the entire graph. The algorithm learns a way to sample and compile information from a node's local neighborhood, aggregating their features and combining it with the node's own features to update its representation, enabling the model to apply the learned aggregated function to generalize for unseen nodes (Hamilton et al., 2017). Another algorithm that leverages on the graph structure of the data is the **Graph Isomorphism Network (GIN)**, designed to match the discriminative power of the Weisfeiler-Lehman graph isomorphism test, a method for graph comparison. It achieves this by using sum-based aggregation of neighbor features combined with a MLP, ensuring the injectivity need to distinguish different neighborhood multisets. This allows the model to learn highly discriminative representations for nodes and graphs (Xu et al., 2019). **Graph Attention Network (GAT)** is another algorithm of this functional group, with a neural network architecture that introduces attention mechanisms to graph-structured data, enabling nodes to weigh the importance of their neighbors when aggregating features, focusing on the most informative ones. The final node representation is computed as a weighted sum of the transformed features of its neighbors, including itself, using the learned attention scores as weights (Velickovic et al., 2018).

In the class of Tree-based lie algorithms that structure their reasoning in a tree-like format, offering interpretability and the ability to handle both categorical and continuous input variables. Its foundation resides in the **Decision Tree (DT)** algorithm, which builds a tree by recursively splitting the dataset based on the feature that yields the greatest information gain, usually measured using entropy, a measure of uncertainty. This approach leads to a hierarchical structure where each decision node represents a condition that partitions the dataset and each leaf regards a class label (Quinlan, 1986). However, a single DT is prone to overfit, especially if it gets too deep. To address this limitation, **Random Forests (RF)**, an ensemble method that shares the same reasoning principle, aggregates the predictions made by multiple DTs, each built on bootstrap samples of the dataset and trained using randomly selected subsets of features, improving generalization and reducing variance, with the final prediction being achieved through majority voting in classification tasks (Breiman, 2001). A gradient boosting approach that also falls into this category of algorithms is the **LightGBM** in which a set of DTs is sequentially trained, each growing fully and using histogram-based binning to accelerate training, with every DT focusing on improving the predictions where the previous one made mistakes. It is optimized to perform well in large datasets, offering fast training with minimal loss in predictive performance (Ke et al., 2017). **XGBoost** is another

boosting ensemble that leverages on DTs, where these grow level-wise and regularization techniques are included to achieve a balance between model complexity and overfitting (Chen et al., 2016).

In Statistical and Probabilistic methods lie algorithms that heavily rely on modelling probabilities and statistical assumptions, grounded in formal statistical theory. **Logistic Regression (LR)** is an example of that, an algorithm used to estimate the probability of a binary outcome using the logit function, which expresses the log-odds of an event as a linear combination of the predictors (Peng et al., 2002). Another would be **Naïve Bayes (NB)**, an algorithm based on Bayes' theorem, assuming conditional independence among features given the class (Lewis, 1998).

Support Vector Machines (SVM) are a category of their own as they stand on considerably different reasoning principles from the other categories seen so far. SVMs are margin-based classifiers, designed to find the optimal separating hyperplane that maximizes the margin between classes. Even though they excel in discriminating between linearly separable classes, they stand out for that ability to distinguish between not linearly separable ones. They do that by introducing slack variables and a regularization term to balance margin and error, handling non-linear patterns by mapping inputs into high-dimensional spaces to discriminate the classes, not possible in the original space (Burges, 1998).

Ensemble methods combine multiple models to improve predictive performance, often outperforming individual learners by enhancing generalization and robustness, as these don't tend to rely on a single model's predictions. **Bagging** methods resort to two principles: bootstrapping, in which multiple random samples of the same size with replacement are extracted and used to train each base classifier, and aggregating, in which the predictions of each base classifier is considered, for classification tasks it usually is by taking majority voting while for regression tasks it typically is an average of the predictions (González et al., 2020). **Boosting** methods build models sequentially and are designed to iteratively improve the predictions made by the previous ones, boosting overall performance (González et al., 2020). **Isolation Forest** is an ensemble method designed for anomaly detection, in which several Isolation Trees are built and recursively partition the dataset based on random feature-value splits. Anomalies, being different from regular instances, are expected to be isolated in fewer splits, resulting in shorter path lengths within these trees. The final anomaly score is obtained by averaging the path lengths across each tree, with lower scores indicating higher likelihood of anomaly (Liu et al., 2008).

Clustering and Instance-Based Methods encompass algorithms that share a relevant computational principle: reliance on similarity-driven decision processes rather than abstract model parameters or distributional assumptions. These algorithms emphasize distance-based reasoning, whether for classification (instance-based learning) or grouping of similar instances (clustering). On the instance based-learning side, one of the most acknowledged algorithm is **K-Nearest Neighbors (KNN)**, mainly for its simplicity and interpretability, assigning class labels

to new instances by taking majority class among the k most similar known instances, or "neighbors", selected based on a distance metric, requiring no training phase nor assuming any underlying probability distribution (Cover et al., 1967). On the clustering side, and intimately connected with the latter, the **K-Means** algorithm partitions a dataset into k clusters by iteratively updating centroids, the center point of each cluster, to minimize variance within each cluster until convergence, once all centroids' position achieve a steady state, although the solution found may not be a global optimum but rather just local as the algorithm is very sensitive to the initial seeded k centroids' positions (Likas et al., 2003). **Hierarchical Clustering** rather forms a nested tree-like structure, called a dendrogram, by iteratively merging or separating clusters based on a linkage criterion, enabling an interpretation of multilevel relationships in the data without predetermining any number of cluster beforehand (Murtagh et al., 2011). The **C-Means** algorithm relaxes the strict assignment to a class constraint posed by other clustering algorithms by allowing for partial membership, introducing fuzzy logic principles, making it ideal to handle uncertainty and model overlapping group boundaries (Askari, 2020).

2.4. RESEARCH SIGNIFICANCE

Despite their pivotal role in ensuring economic stability and sustained growth, FIs are facing growing exposure to sophisticated threats such as fraud and money laundering. These crimes often intersect with fraudulent proceeds urging to be legitimized via obscure laundering schemes (Levi, 2002), designed to circumvent controls and allow criminal to enjoy the ill-gotten money. Traditional rule-based methods are struggling to handle the growing variety and volume of transactional data, leaving FIs way too permeable to this kind of criminal activity (Betron, 2012). ML, on the other hand, can enhance detection systems, especially in a financial transaction environment in which data is characterized by high-dimensionality, class imbalance and scarcity of reliable labeled data (Sudjianto et al., 2010). Yet, compounding these already adversarial conditions is the phenomenon of concept drift, whereby the statistical properties of both fraudulent and non-fraudulent behavior changes over time, driven by the apparently unlimited imagination of criminals (Lu et al., 2019), leading to obsolete static models whose detection performance degrades along, emphasizing the urge for adaptive and data-driven approaches like those ML can offer. This research is significant in both academical and practical terms: it aims to provide state-of-the-art ML approaches to both topics under concept drift while emphasizing the need to work on frameworks that incorporate their intertwined nature instead of handling them independently.

3. METHODOLOGY

3.1. SYSTEMATIC LITERATURE REVIEW

In recent years, money laundering gained some attention in Scientific Research, specifically the blooming application of ML that enables a more accurate approach towards the matter, less prone to human error and capable of understanding non-linear relationships among the data. However, high-entry barriers for researchers are blocking further progress, namely the high-confidential condition of the data that has led to strict legislation on how to use and who can access it. Apart from that, other related topics with which it shares some conceptual and methodological similarities are better established in Scientific Research, namely fraud detection from which knowledge can be extracted and compared to money laundering detection.

With that in mind, an SLR becomes suitable to understand the state-of-the-art of both topics, summarizing the key findings of the application of ML to address those matters and set ground for further research and decision-making. By following a structured framework to develop SLRs like the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) framework (Page et al., 2021), the process ensures transparency and a clear understanding of the steps taken and decisions made along the way to anchor reproducibility of the work done.

3.2. SEARCH PROCESS

To conduct this SLR, a Search Strategy was developed to collect relevant studies on the proposed topics. Several scientific databases were considered but only 2 were selected for their intrinsic characteristics that far distinguish them from any others: Openalex (via openalex R) for its open industry-leading coverage to broaden the search and Scopus for its straightforward approach to emphasize high-impact works. This way, the SLR brings together all the ingredients to achieve a balance between both coverage and relevance while focusing on the most important journals and conferences according to the topics addressed.

An initial search query string was developed to address the proposed Research Question and adapted to the respective databases encompassing keywords related to 4 main pillars:

- **Financial Institution**: the focal agent potentially impacted by the topics;
- **Machine Learning**: the driver of the discussed approaches to handle the topics;
- **Fraud / Money Laundering**: the primary focus of the study;
- **Concept Drift**: the phenomenon of interest that the study aims to investigate under the realm of the topics.

Considering the 2 databases' distinct characteristics, different weights were assigned to the pillars: since Openalex leverages a very large database, an emphasis on concept drift related

keywords was given compared to the others, not only because of it being the phenomenon of particular interest that might differentiate this study from others but also to ensure it was not merely mentioned but rather taken into consideration in the selected works; for Scopus, the same was not applied as it leverages an high-impact oriented database retrieving fewer but high quality works by default.

However, because of the small number of money laundering related works under the mentioned filters in the 2 databases, a second search query string had to be developed to raise more relevant works on that topic, still considering the previously mentioned principles. To that end, “Concept Drift” related keywords were omitted, as this was the pillar cutting off the most works, leading to an extra manual assessment regarding its focus throughout the work while all the others were leveled up in terms of weights. Also, in the “Fraud / Money Laundering” pillar, “Fraud” related keywords were excluded to focus on “Money Laundering” addressing works only.

Finally, to get a clearer vision of what has been achieved in the last decade, only works published between 2014 and 2024 were considered, ensuring an up-to-date SLR. The final search query strings used in each database are depicted in Table 1, alongside the results for each.

3.3. Inclusion and Exclusion Criteria

From the search query strings several works were collected but further refinement was needed to clearly align the chosen ones with this study’s scope. To get an understanding of the application of ML in such topics, only original research works were included in which the authors propose novel frameworks, including techniques and/or algorithms to address those issues, that are assessed using appropriate evaluation metrics afterwards to ensure some degree of scientific quality and rigor. Following the same line of reasoning, peer-reviewed assurance was a must as inclusion criterion and careful considerations were taken regarding the used databases: by default, Scopus deals almost exclusively with peer-reviewed works but Openalex not necessarily due to its open-access nature, which led to the inclusion of a particular filter in the search query strings (“has_oa_accepted_or_published_version:true”) to ensure that criterion was met for journal articles or conference papers, validating their scientific integrity and credibility. Table 3.1 shows the used search queries strings to the respective databases, along with the total number of results retrieved initially from each, followed by the number of those that were excluded and respective motive, and the final number of those that were eventually selected.

Table 3.1 - Queries used in the databases

Database	Query	Total	Excluded	Motive	Results
Openalex (via openalexR)	https://api.openalex.org/works?filter=**default.search**:financial%20institution|financial%20sector|banking%20system|banking%20sector,**default.search**:machine%20learning|artificial%20intelligence|supervised%20learning|unsupervised%20learning|reinforcement%20learning,**default.search**:money%20laundering|terrorist%20financing|financial%20crime|aml%20compliance|fraud%20detection|financial%20fraud,**title_and_abstract.search**:concept%20drift|evolving%20data|dynamic%20data|data%20drift|non-stationary|data%20stationarity|temporal%20dynamics,**has_oa_accepted_or_published_version**:true,**language**:en,**publication_year**:>2013,**publication_year**:<2025	1136	1123 3	Out of scope Books	10
	https://api.openalex.org/works?filter=**title_and_abstract.search**:financial%20institution|financial%20sector|banking%20system|banking%20sector,**title_and_abstract.search**:machine%20learning|artificial%20intelligence|supervised%20learning|unsupervised%20learning|reinforcement%20learning,**title_and_abstract.search**:money%20laundering|terrorist%20financing|financial%20crime|aml%20compliance,**has_oa_accepted_or_published_version**:true,**language**:en,**publication_year**:>2013,**publication_year**:<2025	159	124 30	Out of scope Duplicates	5
Scopus	(**ALL**(("financial institution" OR "financial sector" OR "banking system" OR "banking sector")) AND **ALL**(("machine learning" OR "artificial intelligence" OR "supervised learning" OR "unsupervised learning" OR "reinforcement learning")) AND **ALL**(("money laundering" OR "terrorist financing" OR "financial crime" OR "aml compliance" OR "fraud detection" OR "financial fraud")) AND **ALL**(("concept drift" OR "evolving data" OR "dynamic data" OR "data drift" OR "non-stationary" OR "data stationarity" OR "temporal dynamics"))) AND **PUBYEAR** > 2013 AND **PUBYEAR** < 2025	106	82 12	Out of scope Duplicates	12

(**TITLE-ABS-KEY**(("financial institution" OR "financial sector" OR "banking system" OR "banking sector")) AND **TITLE-ABS-KEY**(("machine learning" OR "artificial intelligence" OR "supervised learning" OR "unsupervised learning" OR "reinforcement learning")) AND **TITLE-ABS-KEY**(("money laundering" OR "terrorist financing" OR "financial crime" OR "aml compliance"))) AND **PUBYEAR** > 2013 AND **PUBYEAR** < 2025	118	98	Out of scope	
		13	Duplicates	6
		1	Books	

Figure 3.1 depicts the flow of works' selection, which started with the retrieval of the results from the search query strings that right after underwent a first validation stage, consisting of manually checking the works' titles, where only around 6% of the initial harvest went on to the following stages. Then, a second validation stage was employed, where the works' abstracts, introduction and conclusion were analyzed to confirm alignment with this study's scope. Finally, only journal articles and conference papers were included, leaving books out of the final selection.

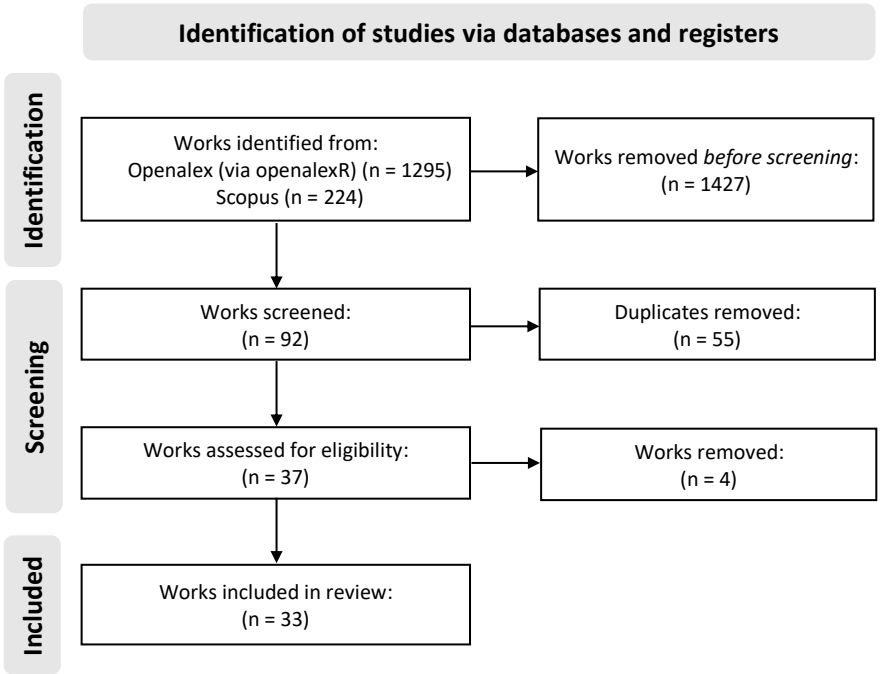

Figure 3.1 - PRISMA information flowchart

4. RESULTS

Table 4.1 displays the results of the SLR, detailed in terms of **Domain**, **Functional Group**, **Algorithms**, **Data Availability** and **Reference**. Each work was associated with the domain it concerns, either money laundering or fraud. Based on the proposed methodology and algorithms employed, each work was assigned with the functional group it shared the most principles with. Additionally, detail on the data available throughout the work was also disclosed. The table is sorted by Domain and then by Reference.

Table 4.1 - Tabulated findings

Domain	Functional Group	Algorithms	Data Availability	Reference
Money Laundering	Support Vector Machines	SVM	Real	(Alkhalili et al., 2021)
Money Laundering	Statistical and Probabilistic	NB	Synthetic	(Caglayan et al., 2022)
Money Laundering	Graph-based	GAT, GIN, GraphSAGE	Real	(Cardoso et al., 2022)
Money Laundering	Neural Networks and Deep Learning	CNN, LSTM, RNN	Real, Synthetic	(Han et al., 2018)
Money Laundering	Statistical and Probabilistic	LR	Real	(Harris et al., 2021)
Money Laundering	Clustering and Instance-based	K-Means, Hierarchical Clustering	Real	(Irshad et al., 2024)
Money Laundering	Neural Networks and Deep Learning	ANFIS	Synthetic	(Jamshidi et al., 2019)
Money Laundering	Tree-based	LightGBM	Synthetic	(Jensen et al., 2023)
Money Laundering	Tree-based	RF	Real	(Ketenci et al., 2021)
Money Laundering	Tree-based	Isolation Forest, RF	Synthetic	(Labanca et al., 2022)
Money Laundering	Clustering and Instance-based	C-Means	Real	(Rocha-Salazar et al., 2021)
Money Laundering	Graph-based	Label Propagation, Isolation Forest	Real	(Tertychnyi et al., 2022)
Money Laundering	Graph-based	GAT	Synthetic	(Yu, 2024)
Money Laundering	Neural Networks and Deep Learning	ANN	Real	(Zhang et al., 2018)
Money Laundering	Graph-based*	LSTM, MLP	Synthetic	(Zhang et al., 2023)

Fraud	Tree-based	RF	Real	(Aftab et al., 2023)
Fraud	Tree-based	RF	Real	(Baabdullah et al., 2024)
Fraud	Neural Networks and Deep Learning	LSTM	Synthetic	(Benchaji et al., 2021)
Fraud	Neural Networks and Deep Learning	ANN	Real	(Chen et al., 2024)
Fraud	Neural Networks and Deep Learning	LSTM, GRU, ANN	Real, Synthetic	(Forough et al., 2021)
Fraud	Tree-based	XGBoost	Real	(Gupta et al., 2024)
Fraud	Clustering and Instance-based	K-Means	Synthetic	(Huang et al., 2024)
Fraud	Neural Networks and Deep Learning	LSTM, NB	Synthetic	(Karn et al., 2022)
Fraud	Ensemble methods	SVM, KNN, RF, Bagging, Boosting	Real	(Khalid et al., 2024)
Fraud	Tree-based	RF	Real	(Muaz et al., 2020)
Fraud	Tree-based	XGBoost	Real	(Noviandy et al., 2023)
Fraud	Ensemble methods	KNN, DT, MLP, LR	Real	(Olowookere et al., 2020)
Fraud	Neural Networks and Deep Learning	ANN	Synthetic	(Sadreddin et al., 2024)
Fraud	Tree-based	RF	Real	(Trivedi et al., 2020)
Fraud	Tree-based	RF, DT	Real	(Velicheti et al., 2023)
Fraud	Graph-based	GAT	Real	(Wang et al., 2024)
Fraud	Neural Networks and Deep Learning	GRU	Real	(Yang et al., 2019)
Fraud	Neural Networks and Deep Learning	CNN, MLP	Real	(Zhu et al., 2024)

*Despite the selected algorithms, the core of the proposed approach is graph-centered.

4.1. QUANTITATIVE ANALYSIS

Despite greater emphasis on the proposed Search Process towards money laundering related works, with the development of an extra search query string for each database, the SLR ended up retrieving more fraud related ones, as seen in Table 4.1, which by itself describes the

distinct current state of research on each domain, with fraud being a more explored field compared to money laundering in the ML spectrum.

Even though the SLR encompasses the period 2014-2024, only works from 2018 onwards made it into this SLR regarding money laundering detection, as seen in Figure 4.1. However, despite a growing trend in the first few years collected, the publishing pace of new original works decreased in the last two years.

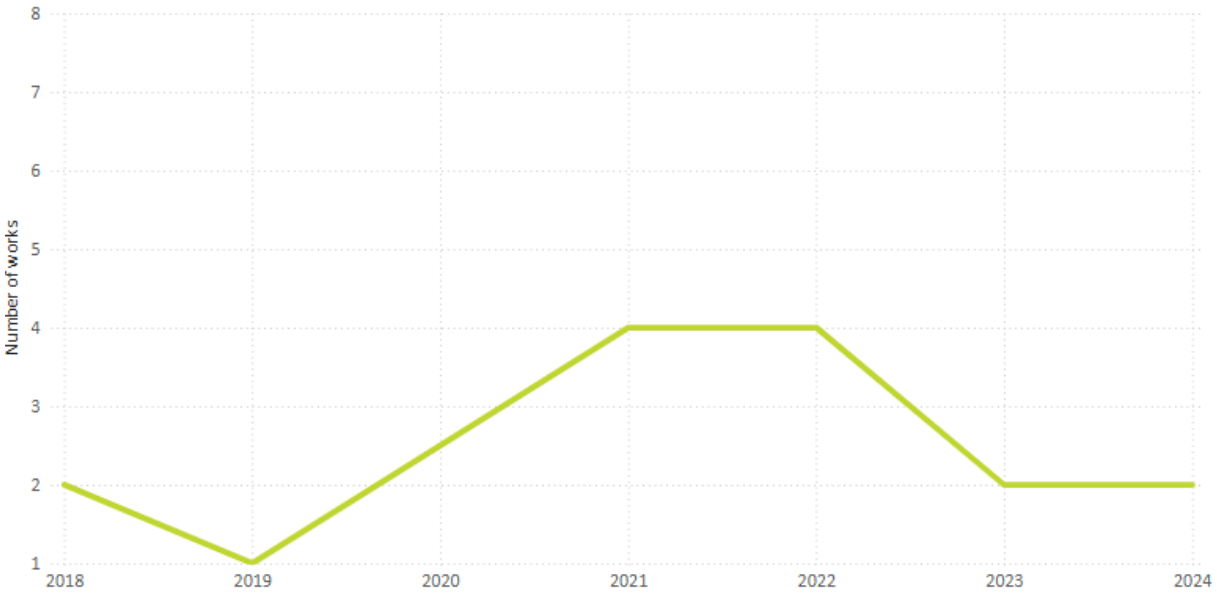

Figure 4.1 - Number of money laundering related works published per year

Similarities and differences are evident for what concerns fraud related works included in this SLR. The fact that only works from the latest years were selected, in this case since 2019, is a clear similarity, which possibly indicates that growing awareness towards the domains in scientific research came around the same time. A key difference, however, is the publishing trend showcased in Figure 4.2, indicating a positively steeper publishing trend in the last couple of years.

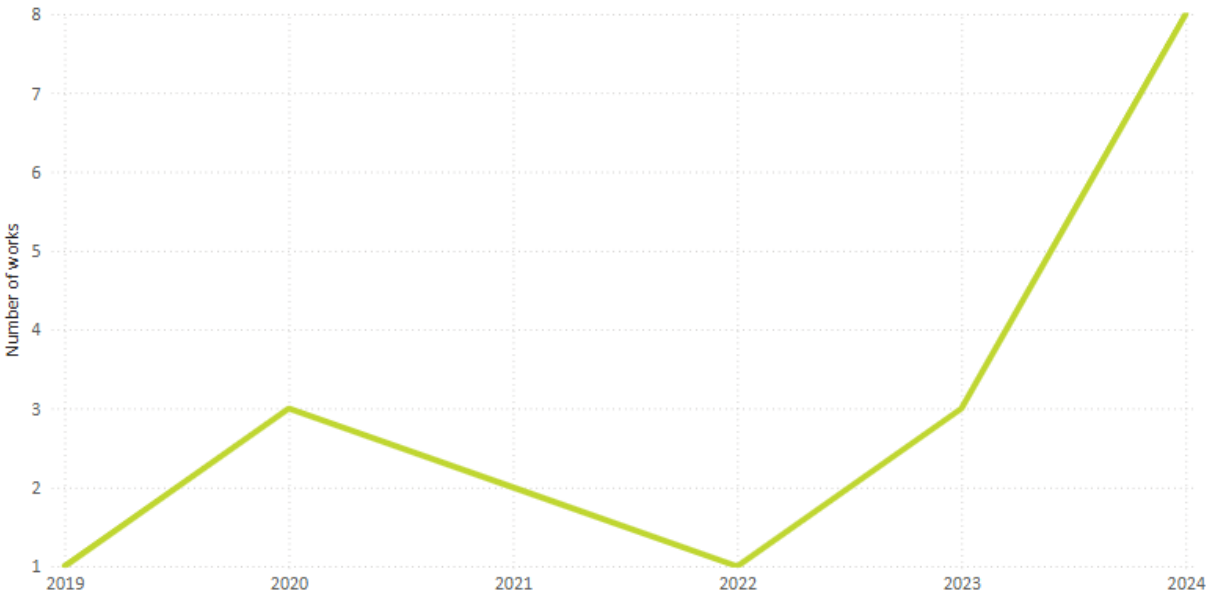

Figure 4.2 - Number of fraud related works published per year

From a technical perspective, the selected money laundering related works are well distributed across several functional groups, according to the principles their methodology is more aligned with. As depicted in Figure 4.3, Graph-based strategies were the most resorted, with Neural Networks and Deep Learning and Tree-based closing the podium in a draw.

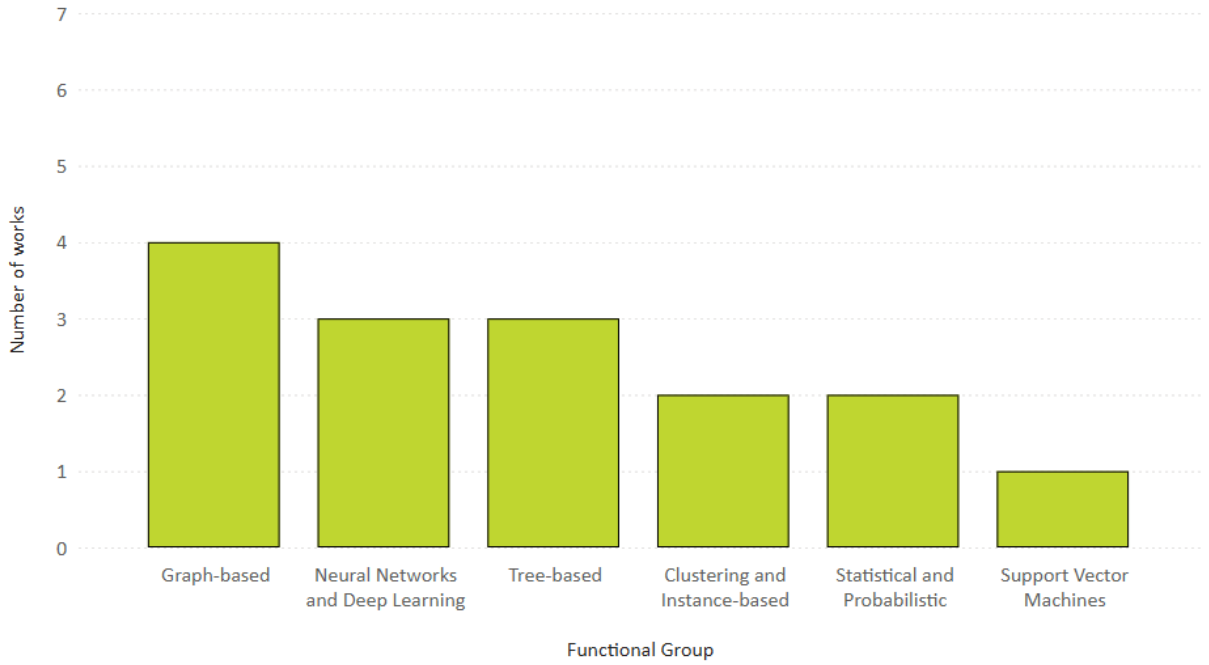

Figure 4.3 - Number of money laundering related works per functional group

Regarding fraud related works, Figure 4.4 displays a much more imbalanced selection of strategies, with the majority of works choosing either Neural Networks and Deep Learning or Tree-based strategies, exactly the second and third most resorted ones in money laundering, indicating potential synergies.

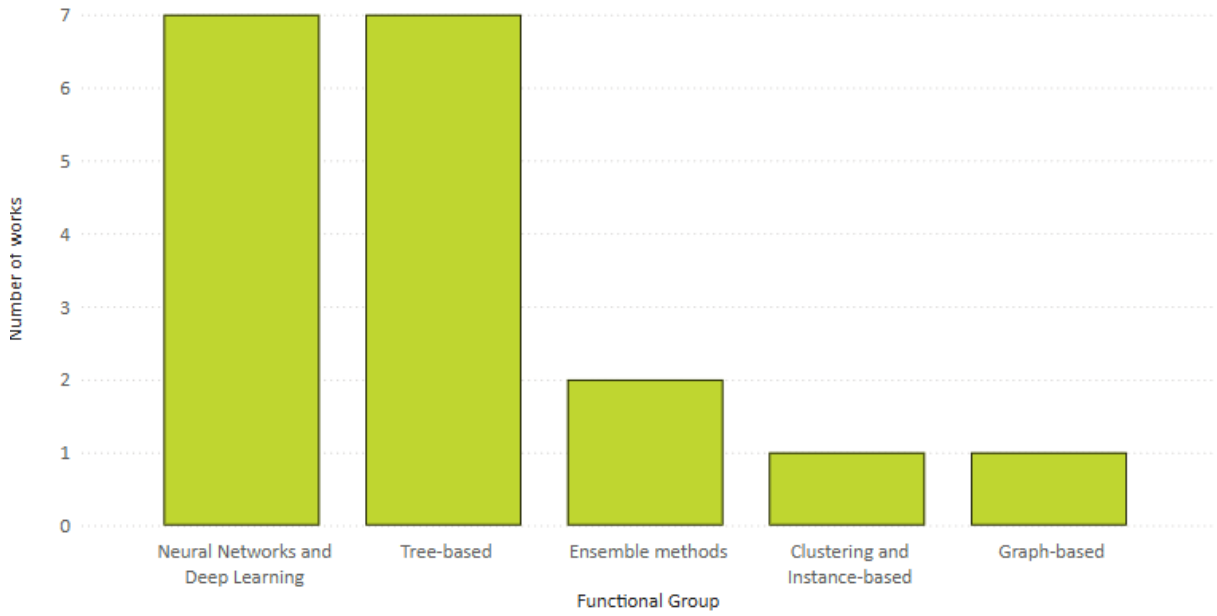

Figure 4.4 - Number of fraud related works per functional group

In terms of data availability during the discourse of each work, money laundering related ones tend to rely more heavily on synthetic grounds compared to its fraud related peers, suggesting that closer partnerships with FIs would allow access to real data and test the proposed methodologies real-world applicability.

Table 4.2 - Data availability in money laundering related works per functional group

Functional Group	Real	Synthetic
Clustering and Instance-based	2	0
Graph-based	2	2
Neural Networks and Deep Learning	2	2
Statistical and Probabilistic	1	1
Support Vector Machines	1	0
Tree-based	1	2

Table 4.3 - Data availability in fraud related works per functional group

Functional Group	Real	Synthetic
Clustering and Instance-based	0	1
Ensemble methods	2	0
Graph-based	1	0
Neural Networks and Deep Learning	4	4
Tree-based	7	0

4.2. QUALITATIVE ANALYSIS

4.2.1. Money Laundering Detection

The network nature of transactional relationships has been widely leveraged by Neural Networks and Deep Learning-based methods across scientific literature. A study employed an ANFIS algorithm to detect customers' risk of involvement in money laundering over a simulated dataset throughout a three-month period, leveraging on the strengths of both ANNs and fuzzy logic to handle complex non-linear problems. Its fuzzy approach was crucial to allocate degrees of partial membership to a fuzzy set of input-output relations of the system, handling the uncertainty and vagueness of the transactional data based on human knowledge, with the ANN component of the algorithm later iterating over those relations until convergence to predict the risk of each customer (Jamshidi et al., 2019). A different study showcasing their potential was developed, in which two distinct sampling schemes, an under-sampling and an over-sampling technique, were applied over a real transactional dataset from a US FI. Several algorithms were trained on an eight-month period and tested on the following five-month period, from which an ANN stood out under both sampling techniques by achieving the highest AUC in both cases out of all algorithms, underlining the suitability to handle rare event detection as time goes by (Zhang et al., 2018). A third study proposed a more sophisticated methodology to take greater advantage of these algorithms. Suspicious transactions originally flagged by a rule-based system were further investigated using sentiment analysis to detect negative sentiment in news and social media towards the involved entities using a CNN. Named entity recognition was performed using a LSTM to identify people and organizations in unstructured text followed by relation extraction employed by a RNN to find connections between those entities. The collected information was used to populate a knowledge graph and a confidence score was assigned to each transaction by aggregating the outputs of each module (Han et al., 2018). Collectively, these studies demonstrate the resilience of this type of algorithms to handle money laundering detection tasks, particularly their ability to handle uncertainty and evolving patterns.

Graph-Based methods have been the most resorted algorithm type used for money laundering detection found by the SLR, as these leverage on the graph-like structure of transactional data. An example of that is a work that employed an LPA to identify communities, groups of tightly connected nodes via edges, over a real dataset from a multinational FI. Communities' graph-based features were extracted and used to refine anomaly detection by an Isolation Forest who identified communities behaving atypically and thus could be involved in group money laundering acts (Tertychnyi et al., 2022). A different work modeled a real-world banking dataset into a bipartite graph of accounts and transactions and explored three different variants of this type of algorithms to learn contextual embeddings for anomaly detection: a GAT, a GIN and a GraphSAGE. The three proved great detection ability displayed by an AUC of at least 90% each underlining this algorithm type's suitability to handle the task, with the GAT variant achieving the best performance (Cardoso

et al., 2022). Another study showcasing the detection suitability of GATs was developed by leveraging synergies between them and a constructed Knowledge Graph, enriching transactional relationships with contextualized edges. A GAT model with a four-layer attention mechanism was trained, resorting to a Focal Loss function to emphasize the cost of misclassified instances, particularly minority class ones. Tested on a synthetic dataset from IBM, the proposed methodology achieved an F1-Score of 71%, demonstrating the strength of attention mechanisms in transactional dataset handling (Yu et al., 2024). One last study resorting to a graph-centric pipeline explored its combination with Deep Neural Networks. The proposed approach focused on modifying the structure of the existing network by adding useful links and removing noisy ones through a graph editor that was guided by a LSTM's feedback, who evaluates graph changes. After editing, a MLP was selected to classify fraudulent transactions over a synthetic dataset (Zhang et al., 2023). All together, these studies demonstrate the growing dominance of Graph-based approaches to handle money laundering detection tasks, particularly their ability to capture complex, evolving fraudulent behavior.

The flexibility and robustness provided by Tree-based methods has also been explored in the realm of money laundering. In a dual-model approach combining unsupervised and supervised techniques, an Isolation Forest was used to identify anomalous behavior within a dataset by iteratively partitioning it and isolating each transaction to assign a respective anomaly score, information to be included in the active learning process of a RF that would classify each transaction as fraudulent or not, with an AML expert revisiting the harder-to-classify ones and flagging them appropriately, providing feedback to be reintroduced in the RF's training exercise. Whilst tested on a synthetic dataset, the proposed approach achieved an AUC of 90%, showcasing reliable detection ability (Labanca et al., 2022). The convenience of RF to handle detection tasks has been proven by yet another study. Leveraging a real transactional dataset from a Turkish bank, the authors applied Fourier Transformation, decomposing it to its constituent frequencies, to obtain a two-dimensional space of time and frequency from which several statistical features were extracted to capture behavioral patterns over time. Combining those with the original dataset and extra CRM data enabled a RF to improve accuracy compared to previous detection systems and reduce false positive alarms (Ketenci et al., 2021). In a different work, a synthetic dataset to benchmark AML detection methods was introduced to provide researchers and FIs with a realistic and privacy-preserving transactional dataset available to the public based on real data from a large Danish bank, mimicking fraudulent changing behavior over time. To validate its utility, the authors selected a set of different algorithms to be trained on the dataset and tested on real transactional data. Results not only proved the dataset's value but also showcased the LightGBM's suitability by consistently obtaining the best results, culminating in the highest AUC among all others (Jensen et al., 2023). Such studies showcase these algorithms' ability to detect fraudulent behavior over time, which can be augmented through auxiliary techniques that emphasize time and frequency.

The heavy reliance on probabilities and statistical assumptions has also proven its value, especially when combined with ML techniques that augment their performance. In a study proposing a graph-based framework, the authors initially converted a synthetic transactional dataset into a graph-like structure, on top of which a feature learning algorithm was employed to generate embeddings that summarized both the graph and relationships between nodes. To address class imbalance, the ADASYN oversampling technique was used, to improve the Gaussian NB's learning ability towards difficult-to-learn instances based on their surrounding neighbors at each moment, synthetically generating new minority ones (Caglayan et al., 2022). A different approach was taken in a study employing an LR algorithm to identify customers at high risk of committing money laundering, addressing the issue as both a binary and multinomial classification problem. The study was conducted using two real datasets from a large bank from Canada, one including customer information and the other regarding monthly transactions spanning over a one-year period. After merging the two, and training the LR among others, it achieved the highest F1-Score, 79%, in both classification problems, (Harris et al., 2021). As shown by the studies, if the right feature learning and sampling techniques are introduced, this kind of algorithms can enhance detection.

SVMs have only been superficially explored in the context of money laundering. In a study that explores its applicability to augment the KYC procedure, Alkhalili et al. proposed a model to check for blocked transactions made by potentially high-risk entities such as sanctioned individuals or politically exposed persons. With the goal of reducing the inherent false-positive rate of the task, the authors trained a SVM with polynomial kernel on historical data to handle non-linear relationships and predict whether incoming transactions should be blocked and later compared to compliance officers' decisions on those same transactions, allowing the model to be refined over time. The study underlines the trend towards phased trust-building, where models progressively gain autonomy in the decision-making process of real FIs on this topic (Alkhalili et al., 2021).

Identifying and isolating similar instances into clusters is one way to approach money laundering detection to raise awareness towards suspicious activity, and Clustering and Instance-Based methods have been effectively applied to do that. A work that combines these algorithms' capabilities with fuzzy logic and an abnormality indicator is that of Rocha-Salazar. In the three-step proposed methodology, compliance experts initially assigned uncertainty-based risk metrics to input variables, feeding a fuzzy logic system to generate probability risk scores; then, clustering-based algorithms were assessed using a compactness measure, from which C-Means stood out; finally, a variance-based abnormality indicator is extracted to refine group comparison. Results demonstrated a reduction in the number of false positives and greater accuracy compared to previous rule-based methods employed in a Mexican FI (Rocha-Salazar et al., 2021). Another application of such algorithms explored its combination with graph network analysis. By leveraging the transactional network dataset from a Pakistani bank, with edge and node level information, it was possible to get a holistic view of the dataset, capturing complex patterns and relationships between accounts. By incorporating

graph-based features with account and transaction level features, a comprehensive feature-set was extracted and served as input for clustering algorithms, K-Means and Hierarchical Clustering, with the latter showcasing greater discriminatory ability, as expressed by a superior AUC (Irshad et al., 2024). Together, these studies underscore the utility of such algorithms to help identify criminal behavior, especially under a scarce label constraint.

4.2.2. Fraud Detection

In the context of fraud, Neural Networks and Deep Learning-Based methods have also been one of the most widely explored, according to the studies identified by the SLR, leveraging on the structure and dynamics of most transactional datasets. Particularly, LSTM algorithms have been successfully applied in this context, as demonstrated by a couple of works found by the SLR. In one, a LSTM was trained with one hidden layer of 15 neurons to analyze temporal relationships among the sequence of transactions carried out by the customers on a simulated dataset based on real transactional data from a Spanish bank, achieving a nearly perfect AUC of 99.5% (Benchaji et al., 2021). Another study highlighting LSTM's detection ability combined it with an error-based bootstrapping technique that re-weighed harder-to-classify transactions in later training iterations. The LSTM worked to find sequential connections between the transactions over time, with its predictions iterating with the bootstrapping technique until the prediction error stopped improving. A NB algorithm was finally used as a meta-classifier of the model to learn from the LSTMs' enhanced predictions, outperforming standard fraud detection approaches (Karn et al., 2022). Its validity to handle the task has been tested in another way, compared to a GRU version of the same proposed methodology. Two ensemble models, each relying on either a set of LSTM or GRU base classifiers, were exposed to and trained on different folds of the dataset, from which predictions were collected and fed as input to a MLP that acted as the meta-classifier of the ensembles. The methodology was tested on two different datasets, a synthetic one and a real transactional dataset from a Brazilian bank, with results demonstrating a slightly more dominant performance from the LSTM ensemble reporting higher average AUC and F1-Score (Forough et al., 2021). In a different approach, a real transactional dataset from a Chinese bank was initially embedded into a graph representation to make connections between accounts clearer which was used to train a GRU algorithm who learned sequential fraud patterns and stored the acquired knowledge in a memory network component at both individual - each account's history of suspicious activity - and group levels - patterns shared across accounts with similar behavior. As new transactions were processed, the memory component of the GRU was updated, allowing for a balance between emerging threats' awareness and previous fraudulent behavior. Results evidenced a strong performance of the model by achieving an AUC of 97% (Yang et al., 2019). In a different investigation resorting to this type of algorithm, Self-Supervised Learning was introduced to understand hidden patterns in transactional data, creating embedding representations of the transactions, that were then enhanced by an embedding-based sampling approach that focused on removing redundant and noisy data points. These served as input of an ANN that provided a likelihood of fraudulent activity for

each embedding, refined by a dynamic threshold that was adjusted in each day of activity available in the datasets to maximize the Recall and minimize the error rate of the model (Chen et al., 2024). In another work exploring the capabilities of ANNs, a synthetic transactional dataset spanning over a six-month period was divided into data chunks, streams of data that respect a fixed time-window (e.g. weekly, monthly). An initial chunk was used to train the first sub-model, an ANN, that learned on that data and preserved its knowledge by freezing the learned weights: that sub-model would then be considered an expert on that data chunk and be stored as an historical sub-model. As new chunks came, new sub-models were trained on the respective chunks and stored. The K most relevant sub-models, according to a Feature Correlation Metric that compared feature vectors from the learned chunk and the current one, would be considered in a dynamic network to provide the final predictions (Sadreddin et al., 2024). In the last work identified by the SLR that emphasizes the capabilities of these algorithms, a MLP, was responsible for selecting the most relevant historical transactions according to the discrepancy between those and the incoming transactions, priorly handled by a reward mechanism employing Deep Reinforcement Learning to maximize the obtained reward, which itself was calculated based on the distance between each historical transaction and the incoming ones, outputting a selection probability distribution vector. Based on the vector, a CNN first extracted the most relevant features from the previously selected training set with an attention mechanism and then output the predicted probabilities for each category. The model was tested on a real transaction dataset from a Chinese bank and results showcased superior performance compared to other peers (Zhu et al., 2024). These studies demonstrate the adaptability and growing sophistication of Neural Networks and Deep Learning-Based methods to address evolving fraudulent patterns.

Although widely used in money laundering detection, fewer Graph-based approaches were identified by the SLR in the context of fraud. A work who opted for such approach was developed using a real-world transactional network with a graph-like structure, connecting multiple nodes via edges. Time encoding operations were employed on those connections to capture their dynamics through relative time differences and the extraction of periodic features. On top of the enriched dataset, a GAT was used to compile and learn information regarding neighboring nodes to detect non-linear fraudulent patterns over time throughout the dataset. Results from the experiment demonstrated good performance, as underlined by an AUC of 79% achieved by the proposed methodology, showcasing great detection potential by these approaches to capture temporal and relational shifts in fraud detection tasks (Wang et al., 2024).

The other most resorted algorithm type in the works identified by the SLR in the context of fraud were Tree-based. All works were conducted on the real, anonymized European Credit Card dataset, containing transactional data spanning over a two-days period. In two of them, comparative studies between RF and other baseline algorithms were conducted after preprocessing the data and applying SMOTE, an over-sampling technique that synthetically generates new samples of the minority class raising awareness towards it during the training

phase. In both cases, the RF was the best performer out of all, achieving the highest discriminatory power (Aftab et al., 2023; Muaz et al., 2020). Other studies explored beyond combinations of RF with over-sampling techniques. In one of those, a novel feedback-based improvement mechanism was introduced to refine the algorithms' detection capability and efficiency by iteratively reintroducing misclassified records into the training data, forcing the algorithm to better understand those cases, with results highlighting an outstanding performance by obtaining an F1-Score of 95%, outperforming all other models (Trivedi et al., 2020). A different approach was considered alongside an implementation of a RF and a DT with the SMOTE technique applied to the transactional dataset, in which a dimensionality reduction technique, PCA, was used to shorten the feature space while conveying the most important information, responsible for most variance, at the cost of interpretability, resulting in great detection accuracy for both algorithms (Velicheti et al., 2023). In a more sophisticated and privacy-preserving concerned implementation of a RF in the context of fraud to align with current legal standards, Federated Learning and blockchain technology were central to the training process. In the study, local learners would be trained independently, representing individual banks, iteratively sharing their parameters with a global learner that would make the final predictions. Out of all the tested local and global learners, RF stood out achieving the highest discriminatory power with an F1-Score of 99%, proving it is possible to perform well while preserving sensitive data (Baabdullah et al., 2024). Two other works resorted to a different Tree-based algorithm, XGBoost. Similar approaches were considered, in which the dataset would be subject to over-sampling techniques to emphasize the presence of the minority class to later be handled by the algorithm who would predict fraudulent behavior. In one, the traditional SMOTE technique was employed, providing a robust framework for effectively detecting fraud (Gupta et al., 2024). In the other, a step-forward was taken by applying a variant of this technique, SMOTE-ENN, that combined SMOTE and Edited Nearest Neighbors to remove noisy synthetic samples by considering their neighbors ensuring that only well-defined ones remained, leading to the highest achieved F1-Score in the study, of 85% (Noviandy et al., 2023). These findings demonstrate the widely usage of Tree-based methods in the context of fraud, explained by their algorithms' robustness and adaptability, but showcases a limitation in what regards dataset testing diversity.

Two other similar studies identified in the SLR focused on showcasing the enhanced detection capability provided by ensembles of a variety of base-classifiers in the context of fraud, both being tested over the anonymized European Credit Card dataset. One of them explored a Stacking ensemble of distinct base-classifiers, DT, MLP and KNN, employing an LR to act as a meta-classifier and learn directly from the predictions of the previous ones. To raise awareness towards the minority class and avoid overfitting, the SMOTE technique was applied to generate synthetic samples of fraudulent transactions during the training of the base-classifiers. Regarding the LR, a cost-sensitive approach was considered, where different misclassification costs were attributed to each class, particularly focused on the amount involved in the respective misclassified transactions, with the results evidencing superior detection performance of the ensemble compared to any sole base-classifier (Olowookere et

al., 2020). In the other study, a voting classifier was preferred, combining the predictions made by several classifiers: SVM, KNN, RF, Bagging of KNNs and a Boosting of RF. The proposed methodology demonstrated great fraud detection ability, showcased by a perfect AUC in the test data (Khalid et al., 2024). These reinforce the enhanced detection performance promoted by the combination of predictions of base-classifiers with distinct reasoning principles in fraud analysis.

Despite its theoretical suitability to handle unlabeled data based on similarity measures, the SLR found little evidence of successful implementations of Clustering and Instance-based methods in the realm of fraud detection. In a study that explored their application, a PCA was employed to reduce dimensionality and keep the most relevant principal components, in the case accounting for 77% of the synthetic transactional dataset's variance. These components were then fed into a K-Means algorithm, with the optimal number of clusters (K = 3) being selected via the elbow method. However, results showcased a limited performance depicted by an F1-Score of 21%, underscoring the limitations of relying solely on unsupervised techniques to handle fraud in FIs (Huang et al., 2024).

4.3. PIPELINE DIAGRAM

Figure 4.5 presents a generic pipeline diagram of a hybrid financial crime detection system, integrating key components relevant to both fraud and money laundering detection identified throughout the SLR. Initially, raw transactional data would be introduced into the system from the FI's source. By integrating these new records with the FI's database, it would be possible to enrich them with available contextual information, such as entity and account metadata, and build graph structures through graph construction, modelling relationships between entities (e.g., customers, accounts) as nodes and interactions (e.g., transactions) as edges. To ensure compliance with data protection requirements, a privacy-preserving layer would be introduced to protect sensitive data (e.g., names, social security numbers) before model training through differential privacy or federated learning, particularly if more than one FI is involved. The GNN layer, exemplified with a GAT algorithm, would be responsible for learning structural representations of the network - embeddings - to capture relational patterns. Those learned embeddings would serve as inputs to a supervised classification model, such as an RF, to assign risk scores or class assignments. The predictions made by the classifier would be subject to an active learning loop in which an expert in the topic would manually label harder-to-classify transactions, reintroducing them into the training process to iteratively improve detection ability, ideal for label-scarce domains such as fraud and money laundering detection. To improve interpretability on the results provided, an explainability layer would be implemented, resorting to XAI techniques like SHAP or Attention Analysis to assign weighted contributions to features in the assignment of a class or risk score, especially for regulatory compliance purposes. Finally, the system would output an alert decision alongside the respective interpretation to be addressed by the appropriate person or team. Additionally, periodic retraining of the model would allow it to gain experience and adapt to fraudulent

emerging trends, leveraging a sliding window strategy in which the training data is segmented into overlapping time-intervals (e.g., 30-day windows with 7-day slides) to allow the model to capture inherent concept drift. This pipeline encapsulates modelling with interpretation and regulation awareness to address hybrid financial crime.

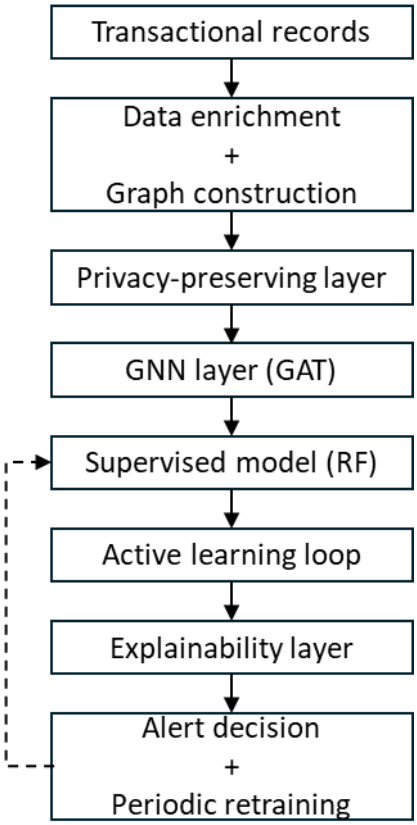

Figure 4.5 - Pipeline diagram for hybrid financial crime detection

5. DISCUSSION

In an ever changing and dynamic environment like that of fraud and money laundering, the handling of concept drift becomes of crucial importance to accurately address both topics as it is the reason why models need to be up to date with the latest trends employed by fraudsters or else they will become obsolete over time. This data phenomenon occurs because fraudsters try to get around every new regulation, making ML models trained on historical data less and less relevant in the long term. With that in mind, researchers must take that into consideration when designing new models to capture deviant behaviors in both domains in an adaptable, sustainable way, particularly in current days of growing awareness and heavier regulation on both domains as FIs are becoming more and more concerned with financial, operational and reputational costs that may arise.

5.1. LINKING FINDINGS TO RESEARCH QUESTIONS

This sub-section will interpret the findings previously presented and discuss them in light of the central research question: **How is machine learning being applied to detect money laundering compared to fraud in financial institutions under concept drift?** By synthesizing the insights from the reviewed literature, this discussion will highlight methodological patterns, contrasts in model usage and the extent to which concept drift is handled across the two domains. The ambition is to move beyond a description of the literature and provide an insightful comparison of how the domains are addressed through ML, particularly in dynamic environments. To that end, three dimensions will be compared across the two domains:

- **Algorithmic Approaches**: captures recurring algorithmic elements found in the literature;
- **Concept Drift Handling**: assess the extent to which the proposed methodologies address evolving patterns;
- **Data Availability**: refers to the accessibility of real datasets to the respective domain.

The first dimension to be assessed is the algorithmic approaches to both domains, according to the identified literature. In fraud detection, there is a clear preference for Tree-based and Neural Networks and Deep Learning strategical approaches. Algorithms like RF, DT and XGBoost dominate across several studies, particularly for their robustness under high class imbalance when combined with SMOTE or SMOTE-ENN and for their suitability to handle tabular datasets like the European Credit Card dataset used to test all their applications in the selected works. Alongside these, Neural Networks and Deep Learning are also widely explored in fraud detection, especially LSTM and GRUs, for their inherent ability to model temporal patterns in sequential data, such as transactional data, to capture evolving fraudulent patterns over time (Benchaji et al., 2021; Forough et al., 2021; Yang et al., 2019). Few exceptions opted not to choose any of these two strategies and still provide relevant insights towards different

fraud tackling approaches. By comparison, money laundering provides a much broader and diverse range of detection strategies. While Tree-based and Neural Networks and Deep Learning algorithmic choices are heavily explored in this context as happens in fraud, evidencing overlapping approaches towards the two domains, there is a slight dominance of Graph-based techniques to money laundering detection. Algorithms like GAT, GIN, GraphSAGE and LPA are leveraged to identify suspicious activity over graph datasets, where nodes are linked via edges, depicting customers/accounts and their activity/transactions with others and providing a more holistic view on the complex, often group-based nature of money laundering (Cardoso et al., 2022; Yu, 2024; Tertychnyi et al., 2022). Additionally, some works provide hybrid strategies, combining Clustering and Instance-based methods with fuzzy logic, ideal to compensate the lack of labels and use domain expert's experience (Rocha-Salazar et al., 2021). All things considered, while there may be overlapping algorithmic approaches to both domains, fraud detection leans towards supervised learning techniques over tabular data mainly, whereas money laundering detection shows a more heterogeneous landscape of approaches, strongly resorting to graph datasets instead to capture the essence of laundering schemes.

Regarding concept drift handling, there are critical differences in the current state approaches across the two domains. In fraud detection, concept drift tends to be handled with greater intentionality and methodological sophistication, as depicted in several works. For instance, dynamic chunk-based modeling (Sadreddin et al., 2024), error-based bootstrapping techniques (Karn et al., 2022) and Deep Reinforcement Learning mechanisms to select the most relevant historical records (Zhu et al., 2024) are great approaches that not only accommodate temporal sequencies but actively aim to adapt to changing data patterns. Moreover, other fraud related works resort to real transactional datasets and perform distinct data splits over time to simulate real-world deployment scenarios and allow the model to adapt to changing data properties (Forough et al., 2021). On the other hand, in money laundering detection, concept drift handling tends to be more implicit, typically resorting to algorithms that have inherent sequence-learning capabilities (Han et al., 2018; Zhang et al., 2023) or employing active learning pipelines during the training process to incorporate the feedback of a domain expert (Labanca et al., 2022). Several other money laundering related works end up relying on static, synthetic datasets and/or not exploring meaningful time horizons during the training of their models, significantly limiting their real-world applicability to dynamic environments, despite recognizing the urge to adapt to evolving laundering schemes. The contrast between the two domains suggests that fraud detection has progressed further in this dimension so far, providing better-structured concept drift handling strategies compared to money laundering's literature that often treats it as a secondary concern, likely due to the added complexity of laundering schemes that make dynamic modeling more challenging.

Data availability is another dimension across which the two domains differ and affect their state-of-the-art. Fraud detection has benefited from a broader availability of real, anonymized

transactional datasets, particularly the European Credit Card dataset used in all identified Tree-based works and a few others. Their availability drives real-world applicability by allowing the proposed methodologies to be subject to a wide range of input structures and patterns and simulate actual deployments through time-based splits as would happen in a real situation. However, fraud related real datasets may lack meaningful transactional timespans (e.g. covering only two consecutive days of transactions), a reason which prompts some researchers to rely on synthetic datasets that cover that gap and allow their models to act and react to data properties changes, like concept drift, that have greater impact over time (Sadreddin et al., 2024; Benchaji et al., 2021). Contrastingly, money laundering related works rely more heavily on synthetic datasets based on pure need, due to the lack of publicly available real datasets. This can be explained by the growing legal, ethical and regulatory barriers that prevent FIs from sharing such sensitive information (Betron, 2012), forcing researchers who cannot get partnerships with FIs to opt for synthetic solutions to train and test their methodologies. Nevertheless, some studies who had access to real datasets, even if anonymized for privacy-preserving purposes, were able to extract meaningful insights regarding their proposed methodologies' real-world applicability. Yet, the lack of real, available datasets for money laundering detection leaves the domain with a urging for benchmarking to allow for a clearer comparison of methodologies and validate their applicability. Overall, fraud detection research is better fulfilled with real available datasets, even if under some degree of anonymization, that provides useful benchmarking for the domain, allowing researchers to choose either one of the available real ones, under their intrinsic limitations, or a synthetic, customized dataset to address any gap considering their limitations in terms of mimicking real-world unpredictability and dynamics, whereas money laundering research do not have such a flexible choice, rather constrained by an evident lack of choices.

Taken all into consideration, the comparison of these three critical dimensions unveils a maturity gap between the domains. Fraud detection benefits from greater real data availability that supports benchmarking and model comparability that propels research, while explicitly handling concept drift. Money laundering remains less consolidated due to unavailability of data which affects standardization and delays real-world applications. Table 5.1 presents a structured summary of the previous comparison, providing practitioners and researchers a clear overview of the maturity of both domains across the discussed dimensions.

Table 5.1 - Comparative table of findings

Domain	Algorithmic Approaches	Concept Drift Handling	Data Availability
Money Laundering	Heterogeneous; Graph-based approaches standout	Mostly implicit by resorting to sequence-learning models or active learning pipelines	Lacks real, open datasets due to legal barriers; heavy reliance on synthetic datasets
Fraud	Dominated by Tree-based and Neural Networks and Deep Learning methods	Mostly explicitly with methods like chunk-based modeling, error-based bootstrapping and temporal splits to simulate deployment	Access to real, anonymized datasets that allow benchmarking; synthetic datasets used to explore evolving patterns

5.2. BROADER IMPLICATIONS

This SLR provides a comprehensive synthesis of the state-of-the-art application of ML to detect both fraud and money laundering in the context of FIs, contributing not only to academic research but also improving Compliance departments' decision-making processes. It allows to get a thorough understanding of the strengths and limitations of the application of particular ML algorithms and techniques according to the specific use case and what is currently being done to tackle the phenomenon of concept drift.

For FIs, this work underscores the need to move beyond static, rule-based systems and opt for more sophisticated models that can adapt to changing criminal strategies over time to ensure their relevance and protect their own clients and operations. The emphasis on concept drift, data imbalance handling and model interpretability reflect today's Compliance and regulatory landscape priorities that need growing additional consideration.

For the research community, the work draws attention to underexplored ML architectures that could enhance fraud and money laundering detection systems if taken full advantage of their intrinsic capabilities. Cross-domain knowledge sharing is also emphasized throughout the work, encouraging researchers to investigate potential synergies between the two topics and also to collaborate closer to FIs in order to get access to real-world transactional data instead of just synthetic datasets that may not resemble the true dynamics behind each environment.

Overall, this work contributes to enhanced decision-making and a more comprehensive view on financial crime detection, balancing innovation with Compliance and regulatory demands, and performance with interpretability and relevance at all times.

5.2.1. Decision Matrix

The decision matrix presented in Table 5.2 serves as a practical tool to guide the practitioners and researchers in the selection of the most appropriate ML strategy in financial crime detection, based on a synthesis of findings of the SLR. It is designed around three key dimensions - **domain** (money laundering vs fraud), **data origin** (real vs synthetic) and **label**

availability (labeled vs unlabeled) – and provides contextualized recommendations for each scenario. Rather than delivering a universal solution, ignoring the inherent differences to each context, the matrix highlights how different combinations of ML algorithms and techniques are appropriate depending on the operational setting. This way, the decision matrix offers a grounded, evidence-based starting point for practitioners and researchers to align a suitable model design to face each scenario.

Table 5.2 - Decision matrix for hybrid financial crime detection strategy

Domain	Data Origin	Label Availability	Recommendation
Money Laundering	Real	Labeled	Combine supervised models (e.g., RF) with GNN components (e.g., GAT) for enhanced context modelling
Money Laundering	Real	Unlabeled	Use unsupervised GNNs (e.g., LPA) and anomaly detectors (e.g., Isolation Forest) to detect fraudulent communities
Money Laundering	Synthetic	Labeled	Blend neural network-based algorithms (e.g., ANFIS, ANN) with graph learners (e.g., GAT) to navigate complex graph structures
Fraud	Real	Labeled	Apply neural network-sequence learners (e.g., LSTM) for temporal modeling, supported by tree-based ensembles (e.g., XGBoost) with SMOTE
Fraud	Real	Unlabeled	Resort to clustering techniques and anomaly detectors (e.g., Isolation Forest), assisted by graph-learners to provide relevant embedding representations (e.g., GAT)
Fraud	Synthetic	Labeled	Leverage tree-based algorithms (e.g., RF) in chunk-based or incremental training setups to identify drifting behavior

5.3. LIMITATIONS

The sole focus on the two mentioned scientific databases, Openalex and Scopus, is one of the main limitations of the SLR as it may have blocked the way to other relevant studies available in other databases only, even though that was priorly addressed through a manual investigation to identify the best-suited scientific databases for the purpose of this work, aligning with both Research Questions and Objectives and covering up the most important conferences and journals to the topic. Moreover, evidence displayed across many of the studies tends to showcase a positive application of ML to address money laundering and fraud detection with an eye on concept drift, which may indicate a publication bias towards favorable outcomes. Another potential bias relates to the selection of words used in the search query strings, which may have narrowed the vision of the SLR towards those cases only, eventually leaving out relevant studies that leveraged on different combinations of keywords

across the 4 searched pillars. Studies developed out of the Scientific Research spectrum, like grey literature, have been left out to ensure some degree of credibility and rigor in the first place but potentially leaving behind relevant findings in those studies. Additionally, the use of synthetic datasets to test the studies' proposed methodologies throughout the SLR is a limitation inherent to the research nature of these sensible topics, from which FIs simply don't allow that information to the public, let alone for free, ultimately conditioning the research community, as a result of more and more restrictive regulations and Compliance programs.

5.4. FUTURE WORK DIRECTIONS

Considering the evidence collected and the limitations of the SLR, it is possible to derive logical follow-up research directions to foster investigation of the topics being analyzed, aiming to i) test and improve robustness under real-world settings, ii) exploring underutilized learning paradigms and iii) aligning performance with growing compliance and privacy requirements. Firstly, although all included studies acknowledge the presence of concept drift in its purest form – changing data patterns over time – there is a clear need for deeper engagement with its different typologies (e.g., virtual, real, sudden, gradual) and how each may affect performance of detection models (Hovakimyan et al., 2024). Future studies should aim to simulate and model these drifts explicitly, enabling a fairer comparison of classifiers under distinct adversarial conditions, preferably based on recurring types of drift in each context – fraud and money laundering. Additionally, drift adaptation strategies must move beyond sole periodic retraining based on fixed intervals or performance degradation triggers: instead, following investigation should explore the incorporation of active learning, in which a human expert in the topic is expected to label harder-to-classify instances to be iteratively reintroduced in the training of the classifier, allowing it to adapt to evolving data patterns based on expert knowledge (Labanca et al., 2022). Secondly, there are promising learning paradigms that remain underexplored in literature such as Reinforcement Learning that have been barely mentioned in a few cases (Zhu et al., 2024) to enhance fraudulent activity detection. Its intrinsic reward-oriented learning capability could leverage historical data and gradually adapt to new behavior, making it highly suitable for dynamic environments like those being studied. Future research on it should explore the applicability of Reinforcement Learning frameworks to the context of fraud and money laundering but considering careful reward design and exploration-exploitation trade-offs. Finally, future work must align detection architectures with growing regulatory and ethical constraints. As emphasis on transparency accelerates, moving towards explainable solutions becomes not only a need but a must and new research should assess the robustness and interpretability of explainability methods such as EBMs and SHAP to detect fraudulent activity under dynamic conditions with the purpose of providing reasonable justifications to stakeholders and raising trust in the decision-making process of these models. Furthermore, the inherent sensitivity of transactional data and customer/account metadata urge a careful privacy-preserving approach due to the growing legal scrutiny around its use. Techniques like differential privacy and federated learning should be explored, serving as enablers of a safer training environment

by introducing controlled noise to protect original data or allowing models to train over decentralized datasets without ever sharing raw data but retaining acquired knowledge (Baabdullah et al., 2024).

To encapsulate all the identified research directions, collaborative investigation strategies should be established to go beyond isolated efforts. It would imply strategic partnerships between FIs, universities and regulatory bodies (e.g., Banco de Portugal, CMVM) to stimulate cooperation between crucial stakeholders within legal sandboxes, controlled environments in which the proposed enhanced architectures can be tested in real and regulated settings, without compromising compliance or privacy requirements. These sandboxes would allow researchers to access representative data and foster trust and transparency between the involved stakeholders, ultimately providing improved detection systems that would benefit economy as whole.

6. CONCLUSION

This study conducted an SLR to investigate the state-of-the-art applications of ML in the context of fraud and money laundering detection under concept drift in FIs. Following the PRISMA framework and drawing from two scientific databases, Openalex (via openalexR) and Scopus, 33 works were selected based on their relevance towards four key pillars: **financial institution**, **machine learning**, **fraud / money laundering**, **concept drift**. The study explored the similarities, differences and intersection between the two topics, analyzing the responsiveness of ML approaches to concept drift and their alignment with real-world settings. Results from the SLR demonstrated that, under the requirements to be selected, fraud detection has been more explored, particularly in later years, with 18 works meeting the inclusion criteria, while money laundering detection has seen a relative slowdown, accounting for 15 works. Key findings denote strong reliance on neural networks and deep learning-based methods in the two topics, suggesting potential for hybrid financial crime investigation strategies. However, there is a heavy reliance on synthetic datasets, which limits the real-world applicability of the proposed detection architectures, let alone the lack of privacy-preserving training mechanisms to protect sensible transactional data and interpretable outputs to provide transparent justifications to stakeholders. Notably, the SLR highlights that the notion of concept drift is underdeveloped in literature, as its different types (e.g., real, virtual, sudden, gradual) have not been specifically acknowledged nor assessed, with many studies resorting to simplistic mitigation strategies such as periodic retraining only. In response, this study proposes several research directions that aim to take full advantage of the application of ML to the topics being discussed, aligning performance with growing compliance and regulatory requirements. By mapping the current landscape and providing directions to close the identified gaps, the SLR contributes to advancing the development of more robust, ethically concerned and applicable financial crime detection systems.

REFERENCES

Abiodun, O. I., Jantan, A., Omolara, A. E., Dada, K. V., Mohamed, N. A., & Arshad, H. (2018). State-of-the-art in artificial neural network applications: A survey. Heliyon, 4, 938. https://doi.org/10.1016/j.heliyon.2018.e00938

Aftab, A. U., Shahzad, I., Sajid, A., Anwar, M., & Anwar, N. (2023). Fraud Detection of Credit Cards Using Supervised Machine Learning Techniques. Pakistan Journal of Emerging Science and Technologies (PJEST), 4(3), 38–51. https://doi.org/10.58619/pjest.v4i3.114

Alkhalili, M., Qutqut, M. H., & Almasalha, F. (2021). Investigation of Applying Machine Learning for Watch-List Filtering in Anti-Money Laundering. IEEE Access, 9, 18481–18496. https://doi.org/10.1109/ACCESS.2021.3052313

Askari, S. (2021). Fuzzy C-Means clustering algorithm for data with unequal cluster sizes and contaminated with noise and outliers: Review and development. Expert Systems with Applications, 165. https://doi.org/10.1016/j.eswa.2020.113856

Baabdullah, T., Alzahrani, A., Rawat, D. B., & Liu, C. (2024). Efficiency of Federated Learning and Blockchain in Preserving Privacy and Enhancing the Performance of Credit Card Fraud Detection (CCFD) Systems. Future Internet, 16(6). https://doi.org/10.3390/fi16060196

Beals, M., DeLiema, M., & Deevy, M. (2015). FRAMEWORK FOR A TAXONOMY OF FRAUD. https://longevity.stanford.edu/financial-fraud-research-center/wp-content/uploads/2016/03/Full-Taxonomy-report.pdf

Benchaji, I., Douzi, S., & el Ouahidi, B. (2021). Credit card fraud detection model based on LSTM recurrent neural networks. Journal of Advances in Information Technology, 12(2), 113–118. https://doi.org/10.12720/jait.12.2.113-118

Betron, M. (2012). The state of anti-fraud and AML measures in the banking industry. Computer Fraud & Security, 5, 5–7. https://doi.org/10.1016/S1361-3723(12)70039-8

Breiman, L. (2001). Random Forests (Vol. 45). https://doi.org/10.1023/A:1010933404324

Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. In Data Mining and Knowledge Discovery (Vol. 2). https://doi.org/10.1023/A:1009715923555

Caglayan, M., & Bahtiyar, S. (2022). Money Laundering Detection with Node2Vec. Gazi University Journal of Science, 35(3), 854–873. https://doi.org/10.35378/gujs.854725

Cardoso, M., Saleiro, P., & Bizarro, P. (2022). LaundroGraph: Self-Supervised Graph Representation Learning for Anti-Money Laundering. Proceedings of the 3rd ACM International Conference on AI in Finance, ICAIF 2022, 130–138. https://doi.org/10.1145/3533271.3561727

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. https://doi.org/10.48550/arXiv.1603.02754

Chen, C. T., Lee, C., Huang, S. H., & Peng, W. C. (2024). Credit Card Fraud Detection via Intelligent Sampling and Self-supervised Learning. ACM Transactions on Intelligent Systems and Technology, 15(2), 1–29. https://doi.org/10.1145/3641283

Cover, T. M., & Hart, P. E. (1967). Nearest Neighbor Pattern Classification. In IEEE TRANSACTIONS ON INFORMATION THEORY (Vol. 24, Issue 1). https://doi.org/10.1109/TIT.1967.1053964

Damásio, B., & Nicolau, J. (2024). Time inhomogeneous multivariate Markov chains: Detecting and testing multiple structural breaks occurring at unknown dates. Chaos, Solitons and Fractals, 180. https://doi.org/10.1016/j.chaos.2024.114478

European Central Bank. (1998). Explanatory notes on statistics on the Monetary Financial Institutions sector. https://www.ecb.europa.eu/stats/pdf/money/mfi/mfi_definitions.pdf??1bb17bb3939ed54de233c530f47853ff

European Central Bank. (2012). THE INTERPLAY OF FINANCIAL INTERMEDIARIES AND ITS IMPACT ON MONETARY ANALYSIS. https://www.ecb.europa.eu/pub/pdf/other/art1_mb201201en_pp59-73en.pdf

Financial Action Task Force. (2012). The FATF Recommendations. www.fatf-gafi.org/en/publications/Fatfrecommendations/Fatf-recommendations.html

Financial Action Task Force. (2021). OPPORTUNITIES AND CHALLENGES OF NEW TECHNOLOGIES FOR AML/CFT. https://www.fatf-gafi.org/en/publications/Digitaltransformation/Opportunities-challenges-new-technologies-for-aml-cft.html

Forough, J., & Momtazi, S. (2021). Ensemble of deep sequential models for credit card fraud detection. Applied Soft Computing, 99. https://doi.org/10.1016/j.asoc.2020.106883

Garza, S. E., & Schaeffer, S. E. (2019). Community detection with the Label Propagation Algorithm: A survey. In Physica A: Statistical Mechanics and its Applications (Vol. 534). Elsevier B.V. https://doi.org/10.1016/j.physa.2019.122058

Goecks, L. S., Korzenowski, A. L., Gonçalves Terra Neto, P., de Souza, D. L., & Mareth, T. (2022). Anti-money laundering and financial fraud detection: A systematic literature review. Intelligent Systems in Accounting, Finance and Management, 29(2), 71–85. https://doi.org/10.1002/isaf.1509

González, S., García, S., del Ser, J., Rokach, L., & Herrera, F. (2020). A practical tutorial on bagging and boosting based ensembles for machine learning: Algorithms, software tools,

performance study, practical perspectives and opportunities. Information Fusion, 64, 205–237. https://doi.org/10.1016/j.inffus.2020.07.007

Goodell, J. W., Kumar, S., Lim, W. M., & Pattnaik, D. (2021). Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. Journal of Behavioral and Experimental Finance, 32. https://doi.org/10.1016/j.jbef.2021.100577

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. Pattern Recognition, 77, 354–377. https://doi.org/10.1016/j.patcog.2017.10.013

Gupta, P., Shukla, S., Kikan, V., & Kumar, A. (2024). Enhancing Fraud Detection in Credit Card Transactions using XGBoost and SMOTE: A Comparative Study. 2024 IEEE International Conference on Smart Power Control and Renewable Energy, ICSPCRE 2024. https://doi.org/10.1109/ICSPCRE62303.2024.10675080

Hamilton, W. L., Ying, R., & Leskovec, J. (2017). Inductive Representation Learning on Large Graphs. https://doi.org/10.48550/arXiv.1706.02216

Han, J., Barman, U., Hayes, J., Du, J., Burgin, E., & Wan, D. (2018). NextGen AML: Distributed Deep Learning based Language Technologies to Augment Anti Money Laundering Investigation. https://doi.org/10.18653/v1/P18-4007

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

Hovakimyan, G., & Bravo, J. M. (2024). Evolving Strategies in Machine Learning: A Systematic Review of Concept Drift Detection. In Information (Switzerland) (Vol. 15, Issue 12). Multidisciplinary Digital Publishing Institute (MDPI). https://doi.org/10.3390/info15120786

Huang, Z., Zheng, H., Li, C., & Che, C. (2024). Application of Machine Learning-Based K-means Clustering for Financial Fraud Detection. In Academic Journal of Science and Technology (Vol. 10, Issue 1). https://doi.org/10.54097/74414c90

Irshad, F., Alkhalifah, T., Alturise, F., & Khan, Y. D. (2024). GCF-MLD: Integrated Approach for Money Laundering Detection using Machine Learning and Graph Network Analysis. IEEE Access, 183961–183972. https://doi.org/10.1109/ACCESS.2024.3510115

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data Clustering: A Review. https://doi.org/10.1145/331499.331504

Jamshidi, M., Gorjiankhanzad, M., Lalbakhsh, A., & Roshani, S. (2019). A Novel Multiobjective Approach for Detecting Money Laundering with a Neuro-Fuzzy Technique. https://doi.org/10.1109/ICNSC.2019.8743234

Jang, J. S. R. (1993). ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Transactions on Systems, Man and Cybernetics, 23(3), 665–685. https://doi.org/10.1109/21.256541

Jensen, R. I. T., Ferwerda, J., Jørgensen, K. S., Jensen, E. R., Borg, M., Krogh, M. P., Jensen, J. B., & Iosifidis, A. (2023). A synthetic data set to benchmark anti-money laundering methods. Scientific Data, 10(1). https://doi.org/10.1038/s41597-023-02569-2

Kaelbling, L. P., Littman, M. L., Moore, A. W., & Hall, S. (1996). Reinforcement Learning: A Survey. In Journal of Artiicial Intelligence Research (Vol. 4). https://doi.org/10.1613/jair.301

Karn, A. L., Ateeq, K., Sengan, S., Indra, G. v., Ravi, L., Sharma, D. K., & Subramaniyaswamy, V. (2022). B-LSTM-NB BASED COMPOSITE SEQUENCE LEARNING MODEL FOR DETECTING FRAUDULENT FINANCIAL ACTIVITIES. Malaysian Journal of Computer Science, 2022(Special Issue 1), 30–49. https://doi.org/10.22452/mjcs.sp2022no1.3

Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T.-Y. (2017). LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Proceedings of the 31st International Conference on Neural Information Processing Systems, 3149–3157. https://doi.org/https://dl.acm.org/doi/10.5555/3294996.3295074

Ketenci, U. G., Kurt, T., Onal, S., Erbil, C., Akturkoglu, S., & Ilhan, H. S. (2021). A Time-Frequency Based Suspicious Activity Detection for Anti-Money Laundering. IEEE Access, 9, 59957–59967. https://doi.org/10.1109/ACCESS.2021.3072114

Khalid, A. R., Owoh, N., Uthmani, O., Ashawa, M., Osamor, J., & Adejoh, J. (2024). Enhancing Credit Card Fraud Detection: An Ensemble Machine Learning Approach. Big Data and Cognitive Computing, 8(1). https://doi.org/10.3390/bdcc8010006

Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. Proceedings of the 2007 Conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in EHealth, HCI, Information Retrieval and Pervasive Technologies, 31, 249–268. https://doi.org/https://dl.acm.org/doi/10.5555/1566770.1566773

Labanca, D., Primerano, L., Markland-Montgomery, M., Polino, M., Carminati, M., & Zanero, S. (2022). Amaretto: An Active Learning Framework for Money Laundering Detection. IEEE Access, 10, 41720–41739. https://doi.org/10.1109/ACCESS.2022.3167699

Levi, M. (2002). Money Laundering and Its Regulation. In Source: The Annals of the American Academy of Political and Social Science (Vol. 582). https://doi.org/10.1177/000271620258200113

Lewis, D. D. (1998). Naive (Bayes) at Forty: The Independence Assumption in Information Retrieval. https://doi.org/https://dl.acm.org/doi/10.5555/645326.649711

Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. In Pattern Recognition (Vol. 36). https://doi.org/10.1016/S0031-3203(02)00060-2

Liu, F. T., Ting, K. M., & Zhou, Z. H. (2008). Isolation forest. Proceedings - IEEE International Conference on Data Mining, ICDM, 413–422. https://doi.org/10.1109/ICDM.2008.17

Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., & Zhang, G. (2019). Learning under Concept Drift: A Review. In IEEE Transactions on Knowledge and Data Engineering (Vol. 31, Issue 12, pp. 2346–2363). IEEE Computer Society. https://doi.org/10.1109/TKDE.2018.2876857

Lyra, M. S., Damásio, B., Pinheiro, F. L., & Bacao, F. (2022). Fraud, corruption, and collusion in public procurement activities, a systematic literature review on data-driven methods. In Applied Network Science (Vol. 7, Issue 1). Springer Science and Business Media Deutschland GmbH. https://doi.org/10.1007/s41109-022-00523-6

Madison, I. (2023). Supply and Demand: The Fundamental Forces of Economics. International Journal of Economics & Management Sciences Perspective, 12(4). https://www.hilarispublisher.com/open-access/supply-and-demand-the-fundamental-forces-of-economics.pdf

Muaz, A., Jayabalan, M., & Thiruchelvam, V. (2020). A Comparison of Data Sampling Techniques for Credit Card Fraud Detection. In IJACSA) International Journal of Advanced Computer Science and Applications (Vol. 11, Issue 6). https://doi.org/10.14569/IJACSA.2020.0110660

Murtagh, F., & Contreras, P. (2011). Algorithms for hierarchical clustering: An overview. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2(1), 86–97. https://doi.org/10.1002/widm.53

Noviandy, T. R., Idroes, G. M., Maulana, A., Hardi, I., Ringga, E. S., & Idroes, R. (2023). Credit Card Fraud Detection for Contemporary Financial Management Using XGBoost-Driven Machine Learning and Data Augmentation Techniques. Indatu Journal of Management and Accounting, 1(1), 29–35. https://doi.org/10.60084/ijma.v1i1.78

Olowookere, T. A., & Adewale, O. S. (2020). A framework for detecting credit card fraud with cost-sensitive meta-learning ensemble approach. Scientific African, 8. https://doi.org/10.1016/j.sciaf.2020.e00464

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., Mcdonald, S., … Mckenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. In The BMJ (Vol. 372). BMJ Publishing Group. https://doi.org/10.1136/bmj.n160

Pal, S. K., & Mitra, S. (1992). Multilayer Perceptron, Fuzzy Sets, and Classification. In IEEE TRANSACTIONS ON NEURAL NETWORKS (Vol. 3, Issue 5). https://doi.org/10.1109/72.159058

Pallathadka, H., Ramirez-Asis, E. H., Loli-Poma, T. P., Kaliyaperumal, K., Ventayen, R. J. M., & Naved, M. (2023). Applications of artificial intelligence in business management, e-commerce and finance. Materials Today: Proceedings, 80, 2610–2613. https://doi.org/10.1016/j.matpr.2021.06.419

Peng, C. Y. J., Lee, K. L., & Ingersoll, G. M. (2002). An introduction to logistic regression analysis and reporting. Journal of Educational Research, 96(1), 3–14. https://doi.org/10.1080/00220670209598786

Quinlan, J. R. (1986). Induction of Decision Trees. In Machine Learning (Vol. 1). https://doi.org/10.1023/A:1022643204877

Rani, M. I. A., Nazri, S. N. F. S. M., & Zolkaflil, S. (2024). A systematic literature review of money mule: its roles, recruitment and awareness. In Journal of Financial Crime (Vol. 31, Issue 2, pp. 347–361). Emerald Publishing. https://doi.org/10.1108/JFC-10-2022-0243

Rocha-Salazar, J. de J., Segovia-Vargas, M. J., & Camacho-Miñano, M. del M. (2021). Money laundering and terrorism financing detection using neural networks and an abnormality indicator. Expert Systems with Applications, 169. https://doi.org/10.1016/j.eswa.2020.114470

Sadreddin, A., & Sadaoui, S. (2024). Chunk-based incremental feature learning for credit-card fraud data stream. Journal of Experimental & Theoretical Artificial Intelligence, 1447–1465. https://doi.org/10.1080/0952813x.2022.2153277

Sanusi, Z. M., Rameli, M. N. F., & Isa, Y. M. (2015). Fraud Schemes in the Banking Institutions: Prevention Measures to Avoid Severe Financial Loss. Procedia Economics and Finance, 28, 107–113. https://doi.org/10.1016/S2212-5671(15)01088-6

Sudjianto, A., Nair, S., Yuan, M., Zhang, A., Kern, D., & Cela-Díaz, F. (2010). Statistical methods for fighting financial crimes. Technometrics, 52(1), 5–19. https://doi.org/10.1198/TECH.2010.07032

Tertychnyi, P., Lindstrom, T., Liu, C., & Dumas, M. (2022). Detecting Group Behavior for Anti-Money Laundering With Incomplete Network Information. Proceedings - 2022 IEEE International Conference on Big Data, Big Data 2022, 2383–2388. https://doi.org/10.1109/BigData55660.2022.10020321

Trivedi, N. K., Simaiya, S., Sharma, S. K., & Lilhore, U. K. (2020). An Efficient Credit Card Fraud Detection Model Based on Machine Learning Methods. International Journal of Advanced Science and Technology, 29(5), 3414–3424. https://www.researchgate.net/publication/341932015

United Nations. (1988). UNITED NATIONS CONVENTION AGAINST ILLICIT TRAFFIC IN NARCOTIC DRUGS AND PSYCHOTROPIC SUBSTANCES. https://www.unodc.org/pdf/convention_1988_en.pdf

United Nations. (2014). Global Programme against Money Laundering, Proceeds of Crime and the Financing of Terrorism (GPML) in the Mekong Region from 2011-2013. https://www.unodc.org/documents/evaluation/Independent_Project_Evaluations/2014/GLOU40_Mekong_Region_Independent_Project_Evaluation_Report_June_2014.pdf

United Nations. (2021). Conference of the Parties to the United Nations Convention against Transnational Organized Crime. https://docs.un.org/en/CTOC/COP/WG.6/2021/2

Velicheti, S. S., Pavan, A. S. H., Tirapathi Reddy, B., Vanisrikala, N., Pranay, R., & Kannaiah, S. K. (2023). The Hustlee Credit Card Fraud Detection using Machine Learning. Proceedings - 7th International Conference on Computing Methodologies and Communication, ICCMC 2023, 139–144. https://doi.org/10.1109/ICCMC56507.2023.10084063

Velickovic, P., Cucurull, G., Casanova, A., Romero, A., Lì, P., & Bengio, Y. (2018). GRAPH ATTENTION NETWORKS. https://doi.org/10.48550/arXiv.1710.10903

Wang, T., & Hsu, C. (2013). Board composition and operational risk events of financial institutions. Journal of Banking and Finance, 37(6), 2042–2051. https://doi.org/10.1016/j.jbankfin.2013.01.027

Wang, Y., Zhan, H., & Jiang, W. (2024). Time Encoding Graph Attention Model for Financial Fraud Detection in Large-scale Financial Social Networks. ACM International Conference Proceeding Series, 70–74. https://doi.org/10.1145/3673277.3673290

Xu, K., Hu, W., Leskovec, J., & Jegelka, S. (2019). HOW POWERFUL ARE GRAPH NEURAL NETWORKS? https://doi.org/10.48550/arXiv.1810.00826

Yang, K., & Xu, W. (2019). FraudMemory: Explainable Memory-Enhanced Sequential Neural Networks for Financial Fraud Detection. Proceedings of the 52nd Hawaii International Conference on System Sciences. https://doi.org/http://hdl.handle.net/10125/59542

Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: Lstm cells and network architectures. In Neural Computation (Vol. 31, Issue 7, pp. 1235–1270). MIT Press Journals. https://doi.org/10.1162/neco_a_01199

Yu, Q. (2024). Enhancing Anti-Money Laundering Systems Using Knowledge Graphs and Graph Neural Networks. Advances in Economics, Management and Political Sciences. https://doi.org/10.54254/2754-1169/2024.18627

Zhang, Y., & Trubey, P. (2018). Machine Learning and Sampling Scheme: An Empirical Study of Money Laundering Detection. Computational Economics, 54(3), 1043–1063. https://doi.org/10.1007/s10614-018-9864-z

Zhang, S., Zhu, Y., & Zhou, D. (2023). TGEditor: Task-Guided Graph Editing for Augmenting Temporal Financial Transaction Networks. ICAIF 2023 - 4th ACM International Conference on AI in Finance, 219–226. https://doi.org/10.1145/3604237.3626883

Zhu, K., Zhang, N., Ding, W., & Jiang, C. (2024). An Adaptive Heterogeneous Credit Card Fraud Detection Model Based on Deep Reinforcement Training Subset Selection. IEEE Transactions on Artificial Intelligence, 5(8), 4026–4041. https://doi.org/10.1109/TAI.2024.3359568

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação

Universidade Nova de Lisboa