Machine Learning Approaches for Money Laundering Detection

Apoorva Dharmendrakumar Patel
Big Data Analytics
Trent University
Peterborough, Canada
appatel@trentu.ca

Raj Rameshbhai Lakkad
Big Data Analytics
Trent University
Peterborough, Canada
rajrameshbhailakkad@trentu.ca

***Abstract*— This study investigates money laundering detection in the worldwide financial system, emphasizing the role of machine learning (ML) technology. ML algorithms offer avenues for upgrading hostile to anti-money laundering (AML) initiatives. Utilizing a dataset from the PaySim test system, which replicates mobile money transactions, the review assesses different ML algorithms' viability in identifying illegal financial activities. Topics covered include feature engineering post-data preparation methods like data cleaning and exploratory data analysis. Through confusion matrix analysis and model comparison, CatBoost, LightGBM, and XGBoost emerge as top performers in fraud detection. These findings feature the significance of robust methodologies and high-level ML algorithms in fortifying AML efforts, especially in digital finance. The research provides an efficient way to deal with utilizing ML techniques for money laundering detection, giving insights and suggestions for reasonable execution in genuine AML systems.**

Keywords— Machine Learning, Money Laundering Detection, Data Transaction

I. INTRODUCTION

The global financial system is being threatened by money laundering. Since criminal organisations are always refining their strategies to take advantage of holes in financial transactions, it is imperative that sophisticated methods be used to identify and stop illicit activities. Machine learning techniques have proven to be groundbreaking in the fight against money laundering because of their ability to process large amounts of data, recognise complex patterns, and adjust to changing conditions. Crime tactics are evolving.

This research is dedicated to exploring how machine learning can significantly enhance our efforts in combating money laundering. The primary objective is to assess the effectiveness of various machine learning algorithms in detecting and preventing money laundering activities. To fortify our fight against money laundering, we are meticulously evaluating a wide range of algorithms, including Random Forest (RF), Decision Tree (DT), K-Nearest Neighbor (KNN), Logistic Regression, XGBoost, Linear Discriminant Analysis, and potentially incorporating some deep learning models.

Money laundering has become easier due to the increasing number of digital transactions and the increasing interconnectivity of financial institutions. Financial institutions and regulatory agencies face tremendous challenges because of criminal organisations using regulatory vulnerabilities and covert operations to shift funds unlawfully. Hence, there is a focused effort to enhance our anti-money laundering (AML) skills and leverage cutting-edge technologies to effectively battle financial crimes.

By harnessing data from transactions, customer information, and timing details, machine learning algorithms can perceive patterns indicative of potential illegal activities. This helps us to identify money laundering instances that warrant more investigation. Notably, this simplified strategy saves critical time and resources by lowering the frequency of false allegations and improving the effectiveness of our anti-money laundering operations.

A thorough assessment of these machine learning algorithms' accuracy in identifying fraudulent transactions is the focus of our research. Furthermore, we are investigating methods to improve these algorithms' interpretability and predictive power. These developments are critical to building resilient models that can adjust to a variety of financial situations, which raises the bar for applying machine learning to the field of money laundering detection.

One of the major advantages of our study is its potential to bolster the security of financial systems through the introduction of sophisticated machine-learning solutions. These solutions not only uncover fraudulent activities but also furnish us with valuable insights into emerging trends and patterns, empowering banks, and regulatory authorities to proactively stay ahead of criminal activities.

Building confidence in financial institutions requires the application of machine learning for anti-money laundering objectives. Fighting financial crimes aggressively will increase the confidence of customers, investors, and other industry players, enhancing the reputation and dependability of the financial sector.

In addition, our study carefully considers the difficulties and constraints involved in using machine learning for money laundering prevention. These problems include maintaining algorithm openness, protecting data privacy, and the requirement for frequent upgrades to keep up with changing money laundering strategies. Overcoming these obstacles is essential to guaranteeing the long-term viability of machine learning-based anti-money laundering systems.

In summary, a notable development in lowering financial crimes and preserving the stability of financial institutions is the use of machine learning in anti-money laundering programmes. We can strengthen our defences against illicit activity by doing ongoing research and improving our machine learning algorithms, which will make the financial world a safer and more secure place overall. Cooperation between banks, regulators, and lawmakers is crucial for using cutting-edge technology and keeping ahead of intricate money laundering schemes, which in turn protects the integrity of global finance.

II. Previous Work

Ongoing innovation and adaptability are essential in the battle against financial crime. Advancements, discoveries, and a diverse range of approaches are what make anti-money laundering (AML) research so vibrant. The main research publications that have influenced our comprehension and use of AML tactics are examined in this section. We will examine how the AML environment is being revolutionised by machine learning, data analytics, and cutting-edge technology.

Our journey begins with a comprehensive survey [1] that offers a bird's-eye view of machine learning (ML) based AML solutions. This study serves as a valuable resource, demystifying the evolving arsenal of techniques employed in detecting and preventing money laundering activities. The authors don't just outline the underlying ML algorithms but also delve into the types of data that fuel these models. Transactional behaviours, customer profiles, and complex relationship networks – all come under scrutiny, revealing the multifaceted approach of ML-powered AML systems.

Next, we encounter a systematic literature review [2] that bridges the gap between AML and financial fraud detection (FFD). Their review highlights a critical yet often overlooked aspect: the interconnected nature of these financial crimes. The authors emphasize the need for robust methodologies and frameworks that can effectively combat both AML and FFD. They offer a thorough perspective for scholars and practitioners alike by classifying research publications according to mixed, qualitative, and quantitative methodologies. This analysis serves as a reminder that while a strong approach to financial crime detection might be harmful, a comprehensive plan is essential for successful mitigation.

Additionally, the interplay of machine learning and sampling schemes in money laundering detection algorithms was investigated using real transaction data from a U.S. financial institution [5]. The study assessed five major machine learning algorithms (Bayes logistic regression, decision tree, random forest, support vector machine, and artificial neural network) and compared two sampling techniques to increase the relative presence of money laundering events in the dataset. Their findings suggested potential advantages of machine learning algorithms in modelling money laundering events, offering valuable insights into the effective utilization of machine learning and sampling schemes in anti-money laundering efforts [5].

Turning around, we look at the research on Saudi Arabia's experience with machine learning for AML [3]. Their study evaluates the effectiveness of supervised machine learning algorithms, with particular attention to Random Forest (RF), Gradient Boosting (GB), Decision Tree (DT), and Nearest Neighbor (KNN). The outcomes are encouraging and demonstrate how well these algorithms can identify enterprises according to risk categories. This study not only contributes to enhancing the detection process but also underscores the importance of machine learning in proactive AML strategies. By leveraging these algorithms, financial institutions can identify high-risk entities before suspicious activities occur.

While machine learning offers a powerful toolkit, the interpretability of its decision-making processes remains a challenge. The application of deep learning (DL) and explainable artificial intelligence (XAI) techniques for detecting money laundering [6]. Their study sheds light on the challenges associated with model interpretability and explainability. The authors call for further research to bridge this gap and integrate DL and XAI techniques effectively into AML efforts. Building trust and maintaining regulatory compliance need deep learning models to be transparent and provide the reasoning behind their judgements, which becomes increasingly difficult as these models get more sophisticated.

A comprehensive survey [7] that serves as a handbook for exploring various machine learning algorithms applicable to AML solutions. They categorize techniques like link analysis, behavioural modelling, risk scoring, and anomaly detection, providing researchers and practitioners with a rich tapestry of options to explore. Their research highlights how important analytics and data preparation strategies are to raising the efficiency of AML systems. To make a masterpiece, a sculptor requires high-quality clay; similarly, solid, and well-prepared data are necessary for ML models to be built.

The fight against money laundering extends beyond financial gain; it also encompasses the financing of terrorism. A survey of machine learning approaches specifically tailored to AML techniques targeting terrorism financing. Their study reviews algorithms that detect money laundering patterns, unusual behaviours, and groups associated with terrorism financing activities [8]. This research highlights the significance of ML-based AML systems not only in safeguarding financial institutions but also in bolstering national security efforts.

A specific aspect of AML – watch-list filtering[9]. A study investigates how machine learning can be applied to automate and streamline this process. They propose incorporating ML components into transaction checking systems, emphasizing the potential for improving accuracy and efficiency in identifying suspicious activities.

In summary, these research papers collectively contribute to advancing our understanding and implementation of machine learning in anti-money laundering strategies, paving the way for more effective and proactive measures in combating financial crimes.

III. Dataset Description

Financial datasets are crucial for research in fraud detection within the realm of mobile money transactions. However, the availability of such datasets, especially in emerging domains like mobile money, is limited due to the sensitive nature of financial data. To address this gap, we present a synthetic dataset generated using the PaySim simulator, which simulates mobile money transactions based on real transaction data extracted from a multinational company's financial logs in an African country. This dataset is scaled down for public use and is designed to aid researchers in developing and evaluating fraud detection methods in the mobile money domain.

With the advancement of mobile money services, financial transactions have seen a transformation, particularly in places where traditional banking institutions are not easily accessible. However, cybercrime targeting mobile money systems is one of the major challenges that this digital revolution has brought us. For these sorts of systems to identify and stop fraud, robust algorithms and processes are

required. High-quality datasets for training and evaluation are another prerequisite for these methodologies.

The dataset introduced in this paper is generated using the PaySim test system, which uses aggregated data from a private dataset obtained from a global organization offering portable mobile financial services across various nations. The original logs provided by the company represent one month of financial transactions in an African country. The PaySim simulator then scales down this dataset to create a synthetic version suitable for public use, while still maintaining the essential characteristics of real-world mobile money transactions.

The synthetic dataset comprises various attributes crucial for studying and analyzing mobile money transactions. These attributes include:

- **Step**: Represents a unit of time in the real world, where one step equals one hour of time. The simulation spans 744 steps, equivalent to 30 days.
- **Type**: Denotes the type of transaction, such as CASH-IN, CASH-OUT, DEBIT, PAYMENT, and TRANSFER.
- **Amount**: Indicates the transaction amount in the local currency.
- **NameOrig**: Refers to the customer initiating the transaction.
- **OldbalanceOrg**: Represents the initial balance of the originator before the transaction.
- **NewbalanceOrig**: This signifies the new balance of the originator after the transaction.
- **NameDest**: Pertains to the recipient of the transaction.
- **OldbalanceDest**: Indicates the initial balance of the recipient before the transaction, excluding merchant accounts.
- **NewbalanceDest**: Represents the new balance of the recipient after the transaction, excluding merchant accounts.
- **isFraud**: Indicates whether the transaction is fraudulent, simulating malicious behaviour aiming to empty funds by transferring to another account and cashing out.
- **isFlaggedFraud**: Flags illegal attempts, such as transferring more than 200,000 in a single transaction, by the business model's control measures.

A. *Key Insights from the Dataset:*

Transaction Patterns: The dataset offers information on a variety of transaction patterns, such as the frequency of various transaction kinds over time (such as CASH-IN, CASH-OUT, etc.). To identify abnormalities or questionable activity, researchers might examine these patterns to comprehend the dynamics of mobile money transactions.

Fraudulent Behaviour: Through the analysis of transactions classified as fraudulent (isFraud = 1), it has been investigated that the traits and tactics employed by fraudulent factors in the virtual setting. As a result, efficient fraud detection algorithms that recognise and stop harmful activity may be developed.

Balance Changes: Although the dataset excludes specific balance-related attributes for fraud detection purposes, it still can be analysed by balance changes (oldbalanceOrg, newbalanceOrig, oldbalanceDest, newbalanceDest) to gain insights into the financial behaviour of customers and recipients involved in legitimate transactions.

Flagged Transactions: The presence of flagged transactions (isFlaggedFraud = 1) highlights attempts to perform illegal activities, such as large fund transfers that violate business control measures. Studying these flagged transactions can contribute to enhancing fraud prevention strategies and improving the security of mobile money services.

Time-Based Analysis: With the dataset structured in hourly steps over 30 days, it has been conducted as a time-based analysis to observe transaction trends, identify peak activity periods, and detect irregularities or patterns associated with fraudulent behaviour at different times of the day or week.

Overall, the dataset offers a comprehensive view of mobile money transactions, fraudulent activities, and control mechanisms, enabling researchers to explore key insights, develop predictive models, and enhance fraud detection techniques in the mobile financial services domain.

IV. METHODOLOGY

Money laundering gives cover for unlawful conduct and threatens the stability of financial institutions globally. Robust anti-money laundering (AML) procedures are necessary for countering this danger effectively. This method outlines a methodical process for looking through financial transaction data to find any signs of possible money laundering. Data preparation, feature engineering, model training, deployment, assessment, and ongoing monitoring and improvement are some of the phases that make up this technique as shown in Fig. 1. The goals of each stage are to improve the money laundering detection systems' efficacy and efficiency.

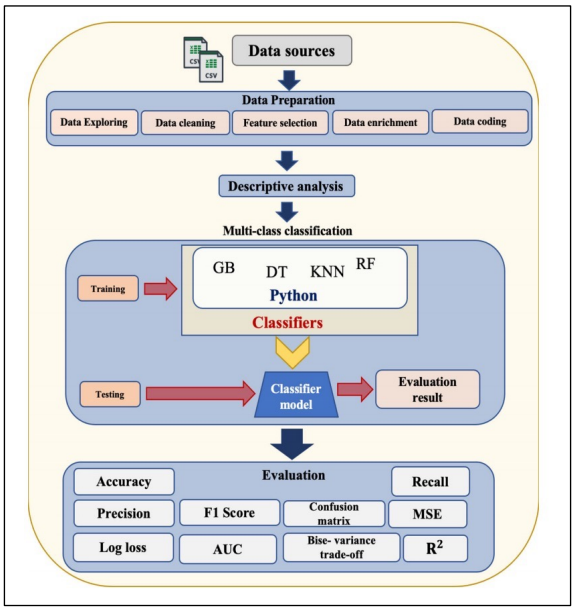

Fig. 1. Phases of Methodology

A. Data Preprocessing

To guarantee the accuracy and dependability of the data utilized for AML analysis, the methodology's initial step is data pre-processing. This phase includes several key steps:

Data Collection: The initial phase of the methodology involves obtaining the financial dataset from an open source. Data has been collected from PaySim test system. The information is obtained from a combined set of data which was collected from a private database belonging to a worldwide organization that provides mobile financial services that can be accessed across multiple countries.

Data Cleaning: Data cleaning requires identifying and fixing issues with data quality, including missing values, duplicate entries, outliers and irregularities. The dataset was ready for analysis using several techniques, including data imputation, data deduplication, and outlier identification.

Exploratory Data Analysis (EDA): To become familiar with the attributes, distributions, and associations between the variables in the dataset, exploratory data analysis (EDA) was carried out. In this stage, patterns, anomalies, and potential insights that could coordinate further examination are found using data visualisation tools like as correlation matrices, scatter plots, and histograms.

B. Feature Engineering

Developing new features or modifying current ones to increase the predictive capacity of machine learning models used for money laundering detection is known as feature engineering, and it is an essential stage in the process. Key steps in feature engineering include:

Transaction Aggregation: We gather transactions by arranging them based on variables like account ID, transaction type, date, and amount. This categorization process, called transaction aggregation, helps us group similar transactions together. We analyze the data to find trends and patterns, then calculate aggregate metrics like total value, frequency, average size, and transaction velocity for meaningful insights.

Derived Features: Derived features are created by extracting additional information from existing data columns. For example, features such as transaction deviation (deviation from typical transaction patterns), suspicious activity flags based on predefined rules, and temporal features (e.g., day of the week, hour of the day) are derived to enrich the dataset and capture relevant information for AML analysis.

Network Analysis: Network analysis techniques are employed to identify complex transaction patterns and relationships between entities such as individuals, businesses, and accounts. Graph-based features such as centrality measures, clustering coefficients, transaction flow metrics, and community detection are used to uncover hidden structures and detect potential money laundering networks.

C. Model Training

The model training phase involves building and training machine learning models to detect potential money laundering activities based on the engineered features. Key steps in model training include:

Model Selection: Selecting the appropriate machine learning algorithm is important to get a successful AML model. Algorithms like Random Forests, Gradient Boosting Machines (GBMs), and Neural Networks are frequently employed to detect money laundering. Many factors impact the technique selection, such as computer resources, interpretability, complexity, and dataset size.

*Feature Scaling and Selection***:** Features are scaled or normalised before model training to guarantee that every variable has an equal impact on the model's performance. To find the most pertinent characteristics that substantially aid in the detection of money laundering operations, feature selection techniques such as principal component analysis (PCA), recursive feature removal, and feature significance ranking are used.

Hyperparameter Tuning: Improving model performance and generalisation requires optimising model hyperparameters. Hyperparameters including learning rates, regularisation strengths, tree depths, and ensemble configurations are fine-tuned using methods like grid search, random search, and Bayesian optimisation to get the best possible model performance.

D. Model Evaluation

The model evaluation phase assesses the performance of the trained AML model using appropriate evaluation metrics and validation techniques. Key steps in model evaluation include:

*Confusion Matrix Analysis***:** The performance of the model is evaluated by comparing the real and predicted labels using a confusion matrix. Metrics like accuracy, precision, recall, F1-score, true positive rate (TPR), false positive rate (FPR), and area under the ROC curve (AUC-ROC) are used to assess how well the model performs in detecting money laundering activities.

Performance Metrics: The model's performance is assessed over a range of criteria using a variety of performance measures. To evaluate the model's trade-offs between true positive rate and false positive rate, precision-recall curves, ROC curves, and area under the precision-recall curve (AUC-PR) are shown, particularly in situations with unbalanced classes characteristic of AML datasets.

In conclusion, the methodology described provides a comprehensive and methodical way to employ network analysis, data analytics, and machine learning for money laundering detection. Among the essential processes that guarantee the reliability and effectiveness of anti-money laundering (AML) systems are feature engineering, model training, assessment, data preparation, continual monitoring, and improvement. Financial crimes necessitate an integrated strategy that takes ethical issues, contemporary technology, cooperative partnerships, and regulatory compliance procedures into account. Thanks to this technique, AML operations are far more successful, and the integrity of the financial industry is protected.

V. RESULTS

A. Descriptive Analysis

Duplicate Check and Data Gaps: The collected datasets were analyzed to gain insights into various aspects of mobile money transactions. All attributes were examined, including ‘Step’, ‘Type’, ‘Amount’, ‘NameOrig’, ‘OldbalanceOrg’, ‘NewbalanceOrig’, ‘NameDest’, ‘OldbalanceDest’, ‘NewbalanceDest’, ‘isFraud’, and ‘isFlaggedFraud’. Notably,

the dataset exhibited no duplicates, ensuring data integrity and no missing values were detected, indicating a complete dataset.

A correlation matrix was constructed to uncover relationships between variables. Variables such as transaction amount ('Amount') and type ('Type') exhibited correlations with fraud indicators ('isFraud').It was observed that fraudulent transactions were primarily associated with 'TRANSFER' and 'CASH_OUT' transaction types.

Fig. 2 shows the relationships between transaction attributes. Strong correlations between transaction amount and fraud indicators are observed.

Histogram plots shed light on transaction amounts, highlighting significantly higher amounts for fraudulent transactions compared to legitimate ones. Furthermore, exploratory analysis revealed transaction patterns over time, with fraudulent transactions exhibiting a consistent distribution, while regular transactions displayed a certain periodicity.

Histogram of transaction amounts for fraud and non-fraud transactions. Fraudulent transactions show a skewed distribution towards higher amounts (see Fig. 3). Analysing transaction patterns according to the day of the week and hour of the day revealed differences in both authentic and fraudulent transactions throughout various periods.

The maximum number of fraudulent transactions on Monday and the minimum number of fraudulent transactions on Thursday (see Fig. 4). The maximum number of non-fraudulent transactions on Saturday and the minimum number of non-fraudulent transactions on Wednesday (see Fig. 5).

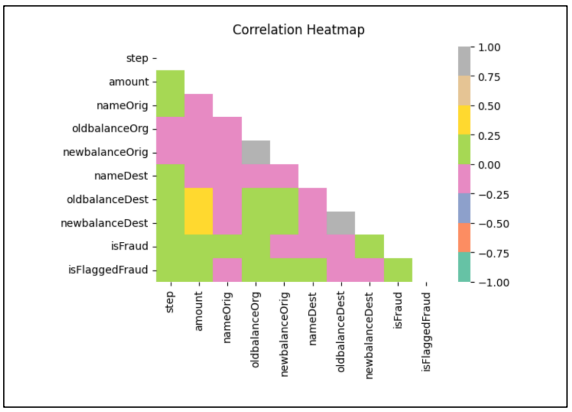

Fig. 2. Correlation heatmap

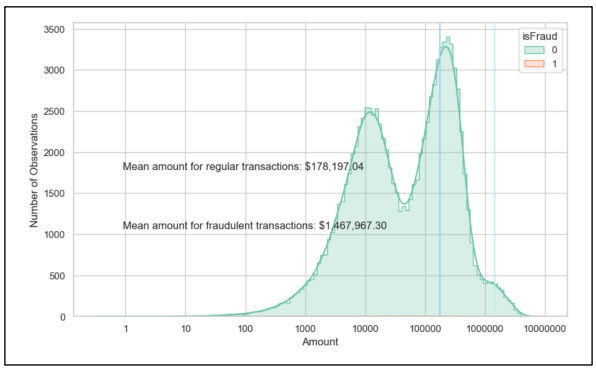

Fig. 3. Histogram of transaction amounts

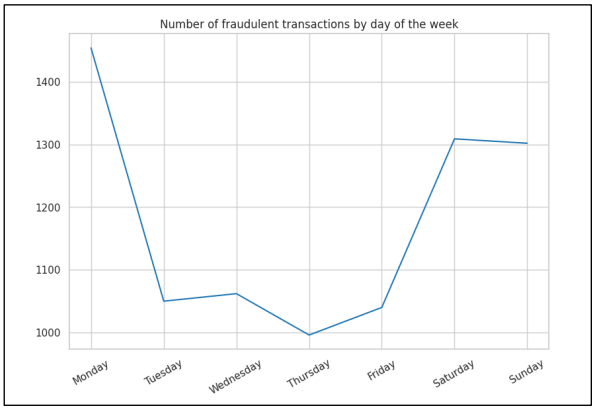

Fig. 4. Fraudulent transaction by day of the week

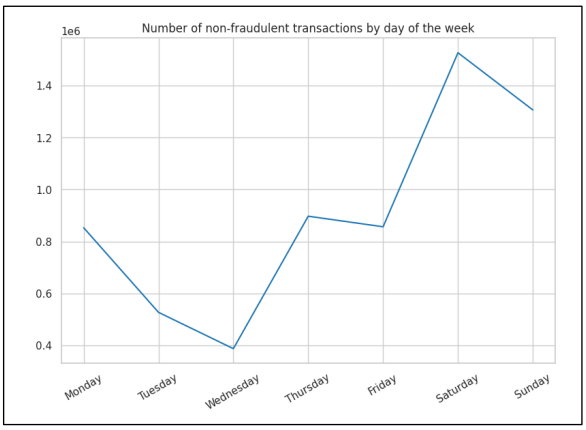

Fig. 5. Non-fraudulent transaction by day of the week

From the EDA it has been analysed that distribution over time of fraudulent trans and those that are not. While the distribution of non-fraudulent transactions is periodic, the fraudulent transaction distribution is steady.

B. Model Evaluation Metrics

To figure out if machine learning models can accurately detect fraud while reducing mistakes like false positives and false negatives, we need to carefully review their performance indicators. This evaluation process includes looking at key metrics and following specific assessment procedures, as outlined in this section.

Confusion Matrix Analysis: A crucial technique for assessing how well categorization models work is the confusion matrix. The model's forecasts are thoroughly broken down, with each one being categorized into four divisions.

- **True Positives (TP):** Instances correctly classified as fraud.
- **True Negatives (TN):** Instances correctly classified as non-fraud.
- **False Positives (FP):** Instances incorrectly classified as fraud.
- **False Negatives (FN):** Instances incorrectly classified as non-fraud.

Each cell in the matrix represents the count of instances classified accordingly, offering insights into the model's classification accuracy and error types.

Confusion matrices were constructed for each machine learning model trained on fraud detection datasets. With the use of these charts, one can assess the model's effectiveness and evaluate how well it distinguished between fraudulent and non-fraudulent transactions. By looking at the numbers for true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), one may determine the model's performance gaps in fraud detection.

For the XGBoost, LightGBM and CatBoost it is the best model because it has high TP and TN rates, while low FP and FN rates.

Performance Metrics: In addition to confusion matrix analysis, several performance metrics were calculated to quantitatively evaluate the effectiveness of fraud detection models. These metrics provided a holistic view of the model's performance across different aspects of classification accuracy.

Precision: The accuracy score shows what percentage of all cases that have been flagged as positive have been correctly recognized as positives.

Recall: Recall quantifies the model's ability to capture all actual positive instances. It measures the proportion of true positives correctly identified by the model among all actual positives. High recall indicates that the model effectively identifies most fraudulent transactions, minimizing false negative errors.

F1-score: The F1-score can be used to assess the degree to which accuracy and memory are correlated. Like a harmony between the two. It is especially useful in scenarios where one class is more common than the other since it takes into account both false positives and false negatives.

Log Loss: A model's ability to forecast probabilistic occurrences is indicated by the log loss. It penalises incorrect classifications by considering the difference between predicted probabilities and actual outcomes. Lower log loss values indicate better predictive accuracy.

Brier Score: The model's probabilistic forecasts are evaluated for accuracy using the Brier score. Lower scores correspond to more accurate probabilistic predictions. It calculates the mean squared difference between anticipated probability and actual outcomes.

AUC-PR quantifies the area under the precision-recall curve, which is particularly useful for imbalanced datasets like fraud detection, where the positive class (fraudulent transactions) is rare compared to the negative class (legitimate transactions). It evaluates the trade-off between precision and recall across different threshold values.

By providing insightful information on many facets of the model's performance, these performance metrics help stakeholders choose and implement the model in a way that is well-informed.

From the model evaluation part, it has been proved that Linear Discriminant Analysis has a low Recall and F1-score, while XGBoost, CatBoost, LightGBM and Decision Tree have high Recall(see Fig. 6). and F1-score (see Fig. 7).

For the Logloss score (see Fig. 8), XGBoost, CatBoost, LightGBM and Decision Tree have low scores which is good, on the other hand, Linear Discriminant Analysis has a high score.

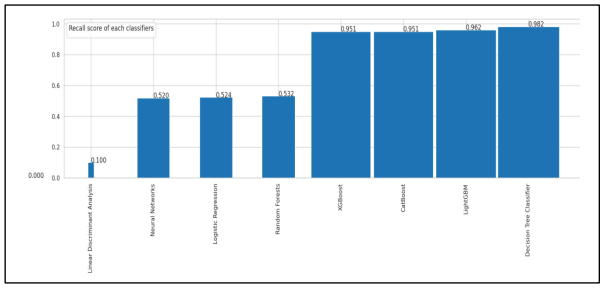

Fig. 6. Recall score for classifier

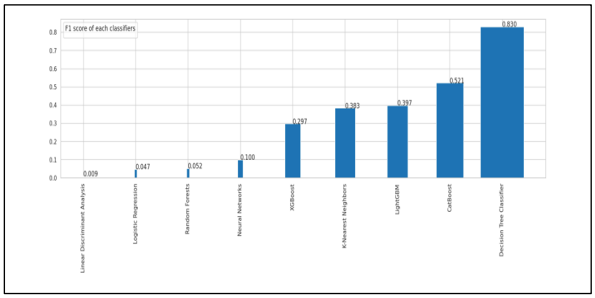

Fig. 7. F1-Score for classifier

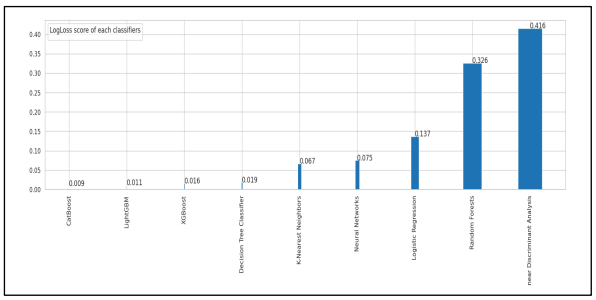

Fig. 8. LogLoss for classifier

C. *Model Comparison with Interpretability*

To identify the top-performing algorithms for fraud detection, performance metrics were compared across all models. Models with higher AUC PR, precision, recall, F1-score, lower log loss, and lower Brier score were prioritised due to their superior performance in accurately classifying fraudulent and non-fraudulent transactions.

In the model building and evaluation part, it has been proven that XGBoost (see Fig. 9)and CatBoost (see Fig. 10) have high AUC PR scores with 0.936 and 0.930 respectively, indicating a strong ability to distinguish between positive and negative cases. Moreover, the Decision Tree classifier also has a good AUC PR score of 0.850, however, we have other better classifiers than this.

Whereas, Linear Discriminant Analysis(see Fig. 11) has a low AUC PR score of 0.204, suggesting it struggles to differentiate between classes.

Ease of understanding and explanation is a crucial factor to consider, especially when it comes to critical domains like fraud detection, even if measuring the efficacy of fraud detection models is vital. While complicated models may perform well, it might be challenging to grasp their decision-making process and spot any biases or mistakes since they are difficult to interpret.

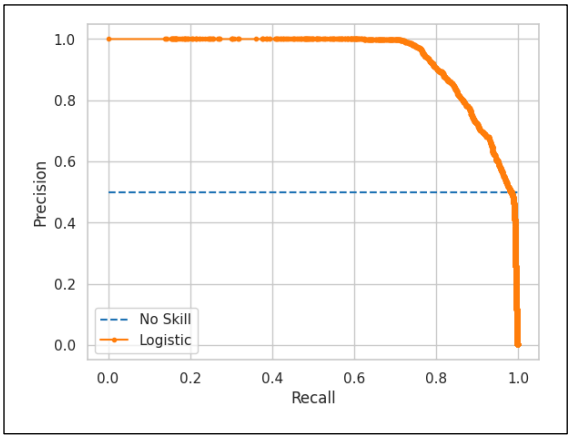

Fig. 9. AUC PR of XGBoost classifier

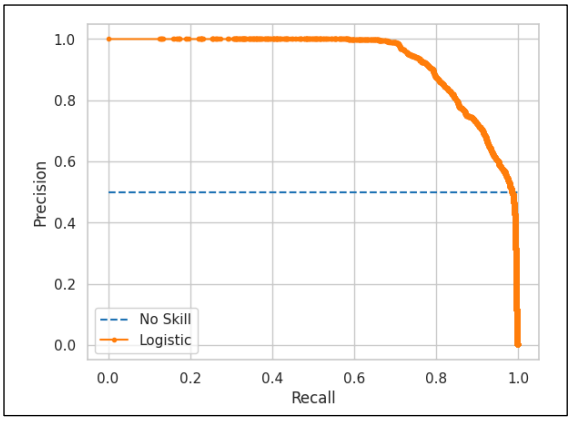

Fig. 10. AUC PR of CatBoost classifier

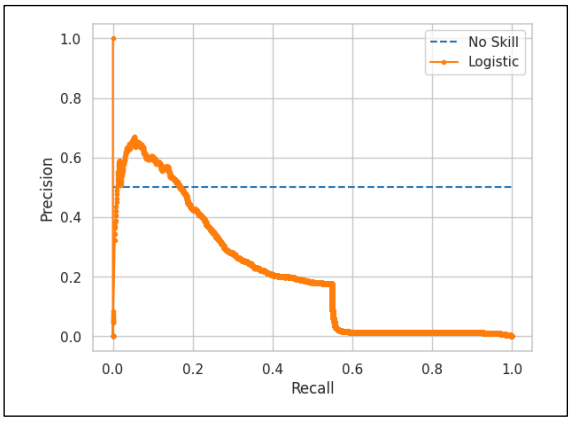

Fig. 11. AUC PR of Linear Discriminant Analysis classifier

VI. Conclusions

The comprehensive evaluation of different ML models for fraud detection uncovered significant experiences in their performance and adequacy. Through examination and comparison, a few key discoveries emerged:

The first one is that, the investigation of model evolution metrics, including accuracy, recall, F1-score, log loss, and Brier score, gave detailed information on each model's predictive capacities. The assessment featured the compromises between various metrics, emphasising the significance of considering numerous factors while choosing an ideal fraud detection model. Notably, models like CatBoost, LightGBM, and XGBoost exhibited robust execution across different metrics, displaying their true potential for fraud detection.

Secondly, the visualization of confusion matrices offered a clear portrayal of the distribution of true positives, true negatives, false positives, and false negatives for each model. These visual representations worked with a better comprehension of the model's classification execution and its capacity to recognize fraudulent transactions while limiting deceptions accurately. The examination of confusion matrices highlighted the significance of adjusting sensitivity and particularity in fraud detection models to accomplish ideal results.

Moreover, the correlation of models through bar plots clarified the general execution of each model across various evaluation metrics. This similar examination enabled to identification of patterns, qualities, and weaknesses across models, helping with the determination of the most appropriate model for deployment.

Moreover, the consideration of AUC PR scores in the model correlation gave a comprehensive assessment of each model's precision-recall trade-off. In general, the evaluation process sheds light on the diverse landscape of fraud detection methodologies, displaying the qualities and limits of different ML models. While specific models exhibited predominant execution in specific measurements, no single model arose as generally prevalent. All things considered, the determination of an ideal fraud detection model ought to be directed by specific use case requirements, functional requirements, and stakeholder inclinations.

Our further exploration and experimentation could focus on refining existing models, investigating ensemble strategies, or utilizing advanced methods like deep learning to figure out how to improve fraud detection abilities. Furthermore, real-time monitoring and evaluation of deployed models will be fundamental to ensure continued viability and versatility in combating developing fraudulent activities. By leveraging the insights acquired from this research, associations can sustain their fraud detection strategies and mitigate financial risks effectively.

VII. References

[1] N. Bakhshinejad, R. Soltani, U. T. Nguyen, and P. Messina, “A survey of machine learning based Anti-Money laundering solutions,” ResearchGate, Oct. 2022, [Online]. Available: https://www.researchgate.net/publication/364326902_A_Survey_of_Machine_Learning_Based_Anti-Money_Laundering_Solutions

[2] L. S. Goecks, A. L. Korzenowski, P. G. T. Neto, D. L. De Souza, and T. Mareth, “Anti - money laundering and financial fraud detection: A systematic literature review,” Intelligent Systems in Accounting, Finance and Management, vol. 29, no. 2, pp. 71–85, Apr. 2022, doi: 10.1002/isaf.1509.

[3] A. A. S. Alsuwailem, E. A. Salem, and A. K. J. Saudagar, “Performance of different machine learning algorithms in detecting financial fraud,” Computational Economics, vol. 62, no. 4, pp. 1631–1667, Sep. 2022, doi: 10.1007/s10614-022-10314-x.

[4] “Synthetic Financial Datasets for Fraud Detection,” Kaggle,Apr. 03, 2017. https://www.kaggle.com/datasets/ealaxi/paysim1

[5] Y. Zhang and P. Trubey, “Machine Learning and Sampling Scheme: An Empirical study of Money Laundering Detection,” Computational Economics (Print), vol. 54, no. 3, pp. 1043–1063, Oct. 2018, doi: 10.1007/s10614-018-9864-z.

[6] D. V. Kute, B. Pradhan, N. Shukla, and A. Alamri, “Deep Learning and Explainable Artificial Intelligence Techniques Applied for Detecting Money Laundering–A Critical Review,” IEEE Access, vol. 9, pp. 82300–82317, Jan. 2021, doi: 10.1109/access.2021.3086230.

[7] Z. Chen, D. Le, E. N. Teoh, A. Nazir, E. K. Karuppiah, and K. S. Lam, “Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: a review,” Knowledge and Information Systems, vol. 57, no. 2, pp. 245–285, Feb. 2018, doi: 10.1007/s10115-017-1144-z.

[8] N. M. Labib, M. A. Rizka, and A. E. M. Shokry, “Survey of Machine Learning Approaches of Anti-money Laundering Techniques to Counter Terrorism Finance,” in Lecture Notes in Networks and Systems, 2020, pp. 73–87. doi: 10.1007/978-981-15-3075-3_5.

[9] M. Alkhalili, M. H. Qutqut, and F. Almasalha, “Investigation of applying machine learning for Watch-List filtering in Anti-Money laundering,” IEEE Access, vol. 9, pp. 18481–18496, Jan. 2021, doi: 10.1109/access.2021.3052313.

[10] E. A. Lopez-Rojas, A. Elmir, and S. Axelsson, “Paysim: a financial mobile money simulator for fraud detection,” Researchgate, pp. 249–255, Sep. 2016, [Online]. Available: https://www.msc-les.org/proceedings/emss/2016/EMSS2016_249.pdf