TELKOMNIKA Telecommunication Computing Electronics and Control
Vol. 22, No. 6, December 2024, pp. 1573~1580
ISSN: 1693-6930, DOI: 10.12928/TELKOMNIKA.v22i6.26163

Utilizing linear regression and random forest models for money laundering identification

Ammar Odeh, Anas Abu Taleb
Department of Computer Science, Faculty of King Hussein School of Computing Sciences, Princess Sumaya University of Technology, Amman, Jordan

Article Info

Article history:

Received Mar 15, 2024
Revised Jul 26, 2024
Accepted Aug 5, 2024

Keywords:

Accuracy
Linear regression
Money laundering
Precision
Random forest
Recall

ABSTRACT

This paper investigates the effectiveness of traditional machine learning techniques, namely linear regression and random forest (RF), in enhancing the detection of money laundering (ML) activities within financial systems. As ML schemes evolve in complexity, traditional rule-based methods struggle with high false favorable rates and a lack of adaptability, prompting the need for more sophisticated analytical approaches. In contrast to the complexities of deep learning models, this study explores the potential of these more accessible machine learning methods in identifying and analyzing suspicious transactional patterns. We apply linear regression and RF models to transactional data to detect anomalous activities that could indicate ML. Our research thoroughly compares these models based on key performance metrics such as accuracy, precision, and recall. The findings suggest that while less complex than deep learning frameworks, linear regression, and RF models offer substantial benefits. They provide a more streamlined, interpretable, and efficient alternative to conventional rule-based systems in the context of ML detection. This study contributes to the ongoing discourse on the application of machine learning in financial crime detection, demonstrating the practicality and effectiveness of these methods in a critical area of financial security.

Corresponding Author:

Ammar Odeh
Department of Computer Science, Faculty of King Hussein School of Computing Sciences
Princess Sumaya University of Technology
Amman 1196, Jordan
Email: a.odeh@psut.edu.jo

1. INTRODUCTION

In the vast and intricate tapestry of global finance, the sinister thread of money laundering (ML) weaves a complex pattern, threatening the very fabric of economic stability and governance [1], [2]. ML, the process of making illegally gained proceeds appear legal, is not just a financial crime but a critical issue that undermines the integrity of financial institutions and the broader economic system. The United Nations Office on drugs and crime estimates that the amount of money laundered globally in one year is 2-5% of global gross domestic product (GDP), amounting to billions of dollars. This staggering figure highlights the urgent need for effective detection and prevention mechanisms [3]-[5].

Detecting ML is a formidable challenge due to the sophistication of the methods used by launderers and the volume of transactions processed by financial institutions. Traditional approaches, such as rule-based systems and threshold settings, are often employed to flag suspicious activities. However, these methods are increasingly proving inadequate [6]-[8]. They are often reactive rather than proactive, heavily reliant on known patterns, and generate a high volume of false positives, leading to inefficiency and investigative fatigue.

Moreover, as money launderers continually refine their tactics, these traditional systems struggle to keep pace, lacking the adaptability and sophistication required to counter such an evolving threat [9]-[11].

In this complex landscape, the emergence of deep learning, a subset of machine learning characterized by its ability to learn from data, represents a beacon of hope [12], [13]. Through their advanced pattern recognition capabilities and adaptability, deep learning algorithms offer a powerful tool against the deception of ML [14], [15]. Unlike traditional methods, deep learning relies not solely on predefined rules or thresholds. Instead, it can analyze vast quantities of data, learning and identifying intricate patterns and anomalies indicative of suspicious behavior. This capability not only enhances the detection of known ML schemes but also paves the way for identifying new, previously unseen tactics [16]-[18]. However, despite the promise shown by deep learning in other domains of fraud detection, its application in anti-money laundering (AML) is still in thenascent stages. Most financial institutions are either in the early phases of adopting these technologies or are hesitant due to the implementation challenges, including the need for large labeled datasets, concerns about explainability, and the complexities of integrating with existing systems [19]-[22].

This study addressed the escalating complexity of ML schemes, which traditional rule-based methods struggle to detect effectively due to high false positive rates and a lack of adaptability. The relevance of this study lies in the critical need to enhance the detection mechanisms within financial systems to combat this pervasive financial crime, which significantly undermines the integrity of financial institutions and the broader economic system. The research employed linear regression and random forest (RF) models, two traditional machine-learning techniques are known for their effectiveness and interpretability in classification tasks. These models were applied to transactional data to detect anomalous activities that could indicate ML. The approach was to utilize these more accessible machine learning methods to identify and analyze suspicious transactional patterns, offering a streamlined and efficient alternative to the complex deep learning frameworks typically employed for such tasks. This research aims to bridge this gap by harnessing the power of deep learning to bolster the fight against ML.

- Development of a tailored deep learning framework: the paper introduces a new deep learning framework specifically designed for the unique requirements of detecting ML in financial transactions.
- Extensive comparison of deep learning models: it provides a detailed comparative analysis of various deep learning models, evaluating their effectiveness in identifying ML activities.
- Implications for key stakeholders: the research discusses the potential impacts of these findings for financial institutions, policymakers, and technology providers, emphasizing how deep learning can revolutionize AML strategies.

2. RELATED WORK

Dalal *et al.* [23] introduces a supervised machine learning model that classifies transactions into three categories: “normal” legal transactions, transactions flagged as suspicious by the bank’s internal systems, and potential ML cases reported to authorities. Uniquely, the model focuses on individual transactions rather than accounts, utilizing XGBoost to achieve an impressive AUC score of 90.7%. The study underscores the importance of including all data, particularly unreported alerts, in the model training process to enhance detection capabilities. Leveraging a comprehensive dataset from Norway, the research conducts a time-validated analysis, ensuring its findings’ practical applicability and relevance. Authentic data confirms that the results reflect real-world scenarios, bolstering the study’s credibility and suggestions for refining machine learning models within AML investigations.

The researchers [16], [24] delves into the alarming prevalence of ML activities, especially in the context of bitcoin transactions. The research illuminates the staggering amounts involved, with an estimated $800 billion to $2 trillion laundered annually, and highlights the significant portion attributed to cryptocurrencies. Recognizing the inadequacies of existing preventive measures, the study advocates for deploying advanced technologies, including deep and machine learning techniques, to develop more effective identification and prevention strategies. Employing the bitcoin elliptic dataset, the research evaluates an array of machine learning models, including deep neural network (DNN), RF, K-nearest neighbors (KNN), and naive Bayes (NB). Among these, DNN and RF emerge as top performers, with RF notably excelling in reducing false positives and achieving a remarkable F1-score of 0.99%. The RF model’s effectiveness is attributed to its adept handling of unbalanced datasets, while the DNN model closely follows with a commendable F1 score of 0.98. The study’s results advocate for the strategic application of sophisticated machine learning algorithms to combat ML in cryptocurrencies, highlighting the importance of minimizing false positives to ensure the practicality and effectiveness of AML policies. The research calls for continuous innovation and collaboration among governments, financial entities, and technology experts to forge robust and adaptive strategies against the ever-evolving landscape of financial crimes.

The central focus is to enhance the precision of AML alert systems through supervised machine learning techniques [25], [26]. The study introduces a triage model as part of the machine learning component, designed to process and score alerts generated by predefined rules. This model allows for the suppression of low-scoring alerts or the prioritization of the alert queue based on these scores, thereby maintaining the system's transparency through rule-based alert generation. Among its notable contributions, the paper introduces the triage model as a novel approach to reduce false positives or prioritize true positives in AML alert processing. It identifies and utilizes specific features to capture entity behavior and relationships in the AML domain, focusing on entity-centric and graph-based characteristics. These include innovative degree-based features and a modified version of GuiltWalker features tailored for delayed labeling scenarios. The paper demonstrates an effective method for graph construction while maintaining memory efficiency for legitimate entities. Evaluated using a real-world banking dataset in an alert suppression scenario, the system shows a significant reduction of 80% in false positives while ensuring over 90% of true positives are detected.

The researchers [27], [28] investigates the application of machine learning techniques in identifying companies involved in ML, with a particular emphasis on data preprocessing and classification methodologies. The study addresses preprocessing techniques, including informative feature selection and categorical feature encoding. It provides an in-depth analysis of classification methods, particularly ensemble machine learning techniques, optimal hyperparameter selection algorithms, and model quality assessment methods. It highlights the importance of AML and countering the financing of terrorism (AML/CFT) measures during the account opening phase for organizations. The research explores integrating various categorical feature transformation techniques and applying cross-validation methods to model this task. The study presents TargetEncoder with double cross-validation as a valuable approach, noting that the gradient-boosting algorithm surpasses decision trees in predictive quality. The trained model is adept at identifying organizations prone to ML and terrorist financing (ML/TF) activities, and the paper discusses the quality of hyperparameter selection, focusing on swiftly identifying optimal sets through Python libraries like hyperopic and optional. The practical application of the study's findings is underscored by developing a Python-based software tool capable of early detection of organizations susceptible to ML, offering insightful recommendations to enhance compliance control procedures.

3. METHOD

Figure 1 outlines the methodo for detecting ML using machine learning. The process begins with data collectionand rigorous data cleaning to ensure data quality. Feature engineering is then performed to extract meaningful indicators from the financial transactions. Data balancing techniques are applied to combat the challenge of imbalanced classes. The method's core lies in the model selection phase, where linear regression and RF models are developed and trained. Finally, the performance of these models is evaluated using metrics such as accuracy, recall, precision, and F-measure to determine their effectiveness in identifying fraudulent transactions. This methodological framework aims to enhance the detection capabilities of financial systems against ML activities.

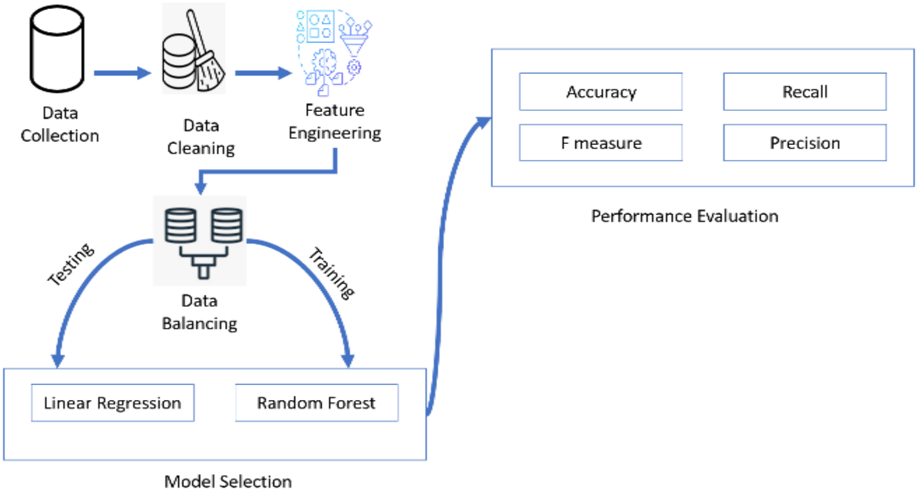


Figure 1. The block diagram for the proposed algorithm

3.1. Data collection and preparation

The initial phase of our method involves collecting transactional data from financial institutions. This dataset includes attributes such as transaction amounts, dates, account types, and other relevant metadata. The primary challenge is the sensitive nature of financial data, requiring strict compliance with data privacy regulations.

3.2. Data cleaning

Data cleaning is crucial to ensure the quality and reliability of the ML models. This step includes handling missing values, correcting errors, and removing duplicates. We employ techniques like imputation for missing data and anomaly detection methods to identify and rectify inconsistencies in the dataset.

3.3. Feature engineering

Feature engineering enhances the ML model's performance by transforming raw data into a more suitable format. We extract meaningful features from transactional data, considering transaction frequency, amount patterns, and historical account behavior. The goal is to encapsulate behavioral patterns that might indicate suspicious activities.

3.4. Data balancing

A significant challenge in fraud detection is the imbalanced nature of datasets, where legitimate transactions vastly outnumber fraudulent ones. To address this, we apply data balancing techniques. For instance, oversampling methods like synthetic minority over-sampling technique (SMOTE) [29] or under-sampling techniques ensure that the models are not biased toward the majority class.

3.5. Model selection

Linear regression and RF models are selected due to their efficacy and clarity in classification roles. Linear regression, while straightforward, offers significant insights into how various features correlate with the probability of a transaction being fraudulent. Conversely, RF, which is a type of ensemble learning, is acclaimed for its precision, capacity to manage voluminous datasets with complex features, and resilience against overfitting. These characteristics make RF particularly valuable for analyzing diverse and intricate data without losing predictive accuracy, thus ensuring robustness in detecting fraudulent activities.

3.6. Model training and validation

The dataset is split into training and testing subsets, ensuring a representative distribution of classes in both. The models are trained on the training set, where they learn to distinguish between legitimate and suspicious transactions. Model validation is performed using cross-validation, which helps assess the model's effectiveness and generalization capability.

3.7. Performance evaluation

The models are evaluated based on several key metrics relevant to anomaly detection tasks, including accuracy, precision, recall, and F1-score. Accuracy measures the overall effectiveness of the model by calculating the percentage of correctly classified instances out of all instances. The F1-score, on the other hand, is the harmonic mean of precision and recall, offering a balanced evaluation when both false positives and false negatives are important. Precision and recall are particularly crucial, as they balance the trade-off between identifying as many suspicious transactions as possible (recall) and maintaining a low rate of false positives (precision).

3.8. Comparative analysis

A comparative analysis of linear regression and RF is conducted to determine their relative strengths and weaknesses in detecting ML. Linear regression is valued for its simplicity and interpretability, making it easier to understand how different variables influence the model's predictions. On the other hand, RF offers higher predictive power by combining multiple decision trees, though it is less interpretable. This analysis also highlights the importance of interpretability in financial applications, where understanding model decisions is crucial for regulatory compliance and trust.

4. RESULTS

This process ensures that the dataset is balanced concerning the class distribution, which is a crucial step in preparing data for machine learning models, especially when dealing with imbalanced datasets that could bias the model's performance. Figure 2 appears to be a bar chart representing the number of transactions categorized as “Fraud” and “Not Fraud.” There are two bars: the blue bar corresponds to “Not Fraud” transactions and the red bar represents “Fraud” transactions. Both bars reach the 1000 mark on the y-axis, indicating the number of transactions. The x-axis is labeled with “Class (0: Not Fraud, 1: Fraud)” to indicate that ‘0’ stands for transactions that are not fraudulent and ‘1’ stands for fraudulent transactions. This chart suggests that the dataset has been balanced to have an equal number of instances in each class, with 1000 transactions for both fraud and not fraud categories. This visualization is useful for quickly understanding the distribution of courses within the dataset after the balancing process described in the code.

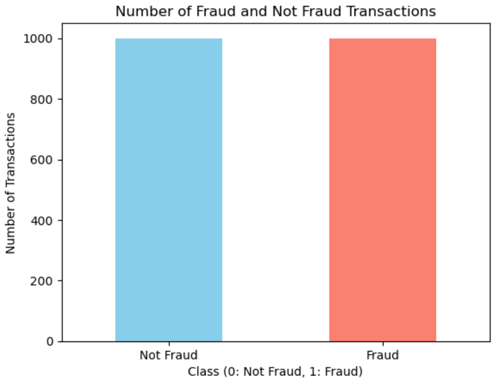
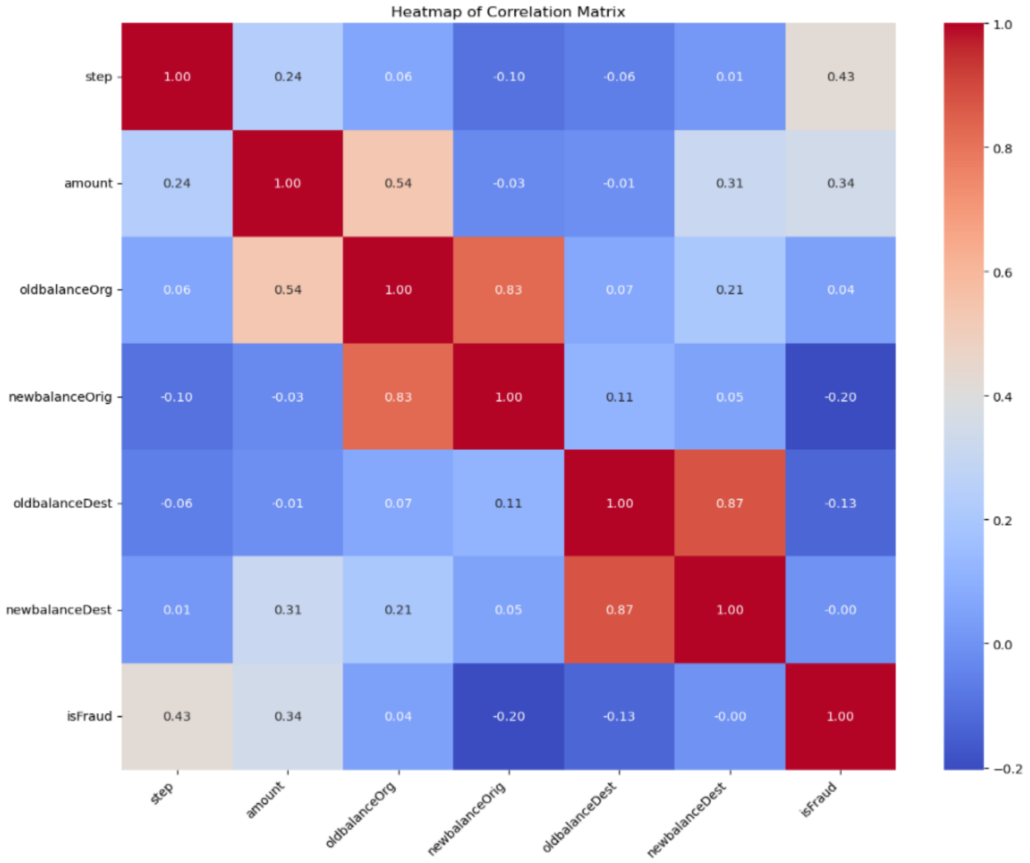


Figure 2. The generated balance dataset

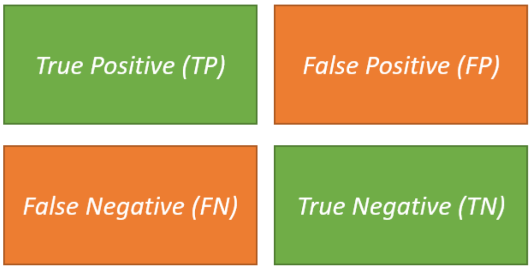
Figure 3 is a heatmap of a correlation matrix, a graphical representation of the correlation coefficients between variables in a dataset. Each square shows the correlation between the variables on each axis. Correlations range from -1 to +1, with +1 meaning a perfect positive correlation, -1 representing a perfect negative correlation, and 0 indicating no correlation.


Heatmap of Correlation Matrix

	step	amount	oldbalanceOrg	newbalanceOrig	oldbalanceDest	newbalanceDest	isFraud
step	1.00	0.24	0.06	-0.10	-0.06	0.01	0.43
amount	0.24	1.00	0.54	-0.03	-0.01	0.31	0.34
oldbalanceOrg	0.06	0.54	1.00	0.83	0.07	0.21	0.04
newbalanceOrig	-0.10	-0.03	0.83	1.00	0.11	0.05	-0.20
oldbalanceDest	-0.06	-0.01	0.07	0.11	1.00	0.87	-0.13
newbalanceDest	0.01	0.31	0.21	0.05	0.87	1.00	-0.00
isFraud	0.43	0.34	0.04	-0.20	-0.13	-0.00	1.00

Figure 3. Heatmap correlation matrix

The colors in the heatmap range from red to blue, with red indicating a higher positive correlation and blue indicating a higher negative correlation. Lighter colors suggest weaker correlations (closer to 0). In determining the importance of features for predicting fraud, those with higher absolute correlation values with "isFraud" could be considered more informative. In this heatmap, "step," "amount," and "newbalanceOrig" seem to be the most significant features of "fraud." However, it's essential to note that correlation does not imply causation and further analysis would be necessary to determine the actual predictive power of these features. Figure 4 provides a visual representation of the four possible outcomes in a binary classification system, commonly used in statistical modeling and machine learning.



True Positive (TP)	False Positive (FP)
False Negative (FN)	True Negative (TN)

Figure 4. Confusion matrix

The measures we used are calculated using (1)-(3).

$$Accuracy = \frac{TP+TN}{(TP+TN+FP+FN)} \tag{1}$$

$$Precision = \frac{TP}{(TP+FP)} \tag{2}$$

$$Recall = \frac{TP}{(TP+FN)} \tag{3}$$

Figure 5 displays a bar chart comparing the performance of two machine learning models, RF, and logistic regression (LR). On a subset of evaluation metrics: accuracy, precision, recall, and F1 score. The metrics are on the x-axis, and the corresponding values are on the y-axis, which ranges from 0 to 1, as these metrics are typically between these values.

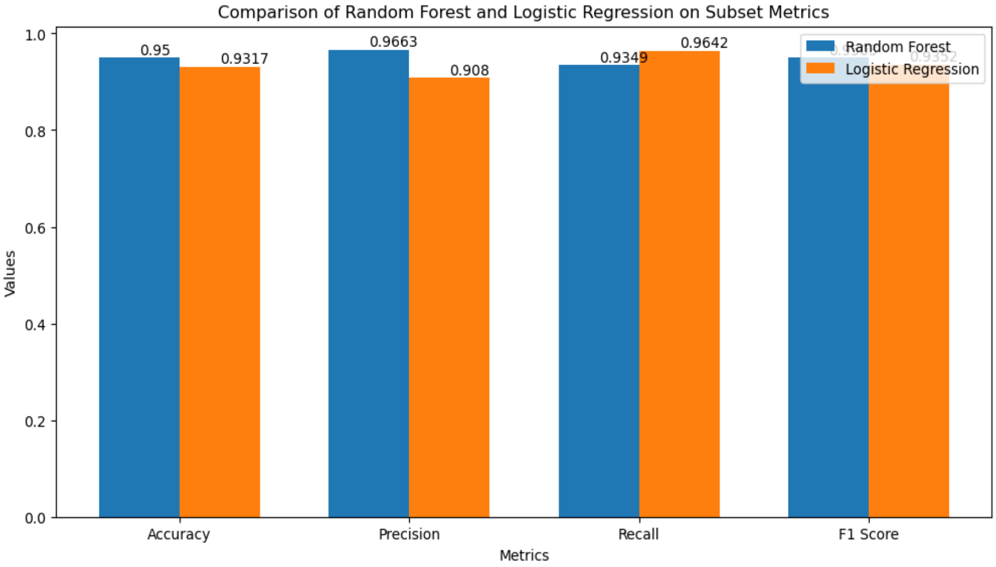


Figure 5. Compare between LR and RF

Each metric has two adjacent bars, one for RF (in blue) and one for LR (in orange), allowing for an easy comparison between the two models on each metric. The values of the metrics are displayed the respective bars, indicating the performance scores for each model.

- Accuracy: the RF model has an accuracy of approximately 0.95, while the LR model has an accuracy of roughly 0.9317.
- Precision: RF shows a precision of approximately 0.9663, and LR offers an accuracy of around 0.908.
- Recall: RF has a recall of about 0.9349, slightly less than its precision, while LR has a recall of approximately 0.9642, higher than its precision.
- F1 score: the F1 score for RF is approximately 0.95, and for LR, it's roughly 0.932.

The bar chart effectively illustrates that the RF model outperforms the LR model in accuracy, precision, and F1 score, while the LR model has a slightly higher recall score than the RF model. This visualization helps assess which model performs better according to each specific metric. The study's findings revealed that while linear regression and RF models are less complex than deep learning frameworks, they still offer substantial benefits in the detection of ML. The comparative analysis of these models based on key performance metrics such as accuracy, precision, and recall demonstrated their practical effectiveness in identifying.

5. CONCLUSION

This paper has presented a comprehensive analysis of the application of linear regression and RF models in identifying ML activities. Through methodical data preprocessing, feature selection, and model training, we have demonstrated that machine learning can significantly enhance the detection of fraudulent transactions in large datasets. Our results indicate that the RF model outperforms linear regression in terms of accuracy, precision, and F1 score, suggesting that the former is better suited for this classification task. The robust nature of RF in handling the non-linear and complex patterns typically associated with ML activities accounts for its superior performance.

Furthermore, the recall metric highlighted the sensitivity of the LR model, underscoring its potential utility in scenarios where capturing as many fraudulent transactions as possible is critical, even at the expense of increasing false positives. This research significantly contributes to the ongoing discourse on the application of machine learning in financial crime detection. It demonstrates the practicality and effectiveness of linear regression and RF models as potent tools for enhancing the detection capabilities of financial systems against ML activities, thus providing a valuable resource for financial institutions aiming to bolster their defenses against such sophisticated crimes.

REFERENCES

[1] S. Shukla, K. Bisht, K. Tiwari, and S. Bashir, "Comparative Study of the Global Data Economy," in *Data Economy in the Digital Age*, ed: Springer, 2023, pp. 63-86, doi: 10.1007/978-981-99-7677-5_4.

[2] T. S. Sobh, "An intelligent and secure framework for anti-money laundering," *Journal of Applied Security Research*, vol. 15, pp. 517-546, 2020, doi: 10.1080/19361610.2020.1812994

[3] R. S. Fullerton and M. J. Patterson, *Murder in our midst: comparing crime coverage ethics in an age of globalized news*: Oxford University Press, USA, 2021.

[4] Q. Abu Al-Haija, A. Odeh, and H. Qattous, "PDF Malware Detection Based on Optimizable Decision Trees," *Electronics*, vol. 11, p. 3142, 2022, doi: 10.3390/electronics11193142.

[5] W. N. W. Ahmad, R. A. Hassan, M. Zakaria, S. N. Md Nazir, and D. Suhartini, "Predicting the adoption of "Buku Kedai": a digital bookkeeping application among Malaysian micro-entrepreneur," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 22, no. 1, pp. 95-103, Feb. 2024, doi: 10.12928/telkomnika.v22i1.25607.

[6] M. N. M. Yunoh, M. Muhamad, Z. K. C. Musa, and N. Bahari, "Understanding the factors influencing the adoption of e-wallets by Malaysian youth," *TELKOMNIKA (Telecommunication Computing Electronics and Control),* vol. 21, no. 6, pp. 1298-1307, Dec. 2023, doi: 10.12928/telkomnika.v21i6.24082.

[7] R. Saxena, D. Arora, and V. Nagar, "Supervised Machine Learning Framework for Deanonymization of Transactional Entities in Decentralized Infrastructure Environment," *Research Square Preprint*, 2023, doi: 10.21203/rs.3.rs-3734345/v1.

[8] A. Tedyyana, O. Ghazali, T. Asnafi, O. W. Purbo, N. Z. Harun, and F. Riza, "Transforming the voting process integrating blockchain into e-voting for enhanced transparency and securiy," *TELKOMNIKA (Telecommunication Computing Electronics and Control),* vol. 22, no. 2, pp. 311-320, 2024, doi: 10.12928/telkomnika.v22i2.25758.

[9] X. Sun *et al.*, "MonLAD: Money laundering agents detection in transaction streams," in *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, 2022, pp. 976-986, doi: 10.1145/3488560.3498418.

[10] J.-de-J. Rocha-Salazar, M.-J. Segovia-Vargas, and M.-del-M. Camacho-Miñano "Money laundering and terrorism financing detection using neural networks and an abnormality indicator," *Expert Systems with Applications,* vol. 169, p. 114470, 2021, doi: 10.1016/j.eswa.2020.114470.

[11] B. Youssef, F. Bouchra, and O. Brahim, "State of the Art Literature on Anti-money Laundering Using Machine Learning and Deep Learning Techniques," in *The International Conference on Artificial Intelligence and Computer Vision*, 2023, pp. 77-90, doi: /10.1007/978-3-031-27762-7_8.

[12] A. Abdollahi, S. B. Arzandeh, and M. Sheibani, "Privacy and safety of narrowband internet of things devices," *TELKOMNIKA (Telecommunication Computing Electronics and Control),* vol. 22, no. 4, pp. 969-975, 2024, doi: 10.12928/telkomnika.v22i4.26279.

[13] R. Kartadie, F. Erwis, A. F. Boy, and A. H. Nasyuha, “A combination of hill cipher and RC4 methods for text security,” *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 22, no. 2, pp. 351-361, 2024, doi: 10.12928/telkomnika.v22i2.25628.

[14] A. A. S. Alsuwailem and A. K. J. Saudagar, “Anti-money laundering systems: A systematic literature review,” *Journal of Money Laundering Control*, vol. 23, pp. 833-848, 2020, doi: 10.1108/JMLC-02-2020-0018.

[15] N. A. Al-Suwaidi and H. Nobanee, “Anti-money laundering and anti-terrorism financing: a survey of the existing literature and a future research agenda,” *Journal of Money Laundering Control*, vol. 24, no. 2, pp. 396-426, 2021, doi: 10.1108/JMLC-03-2020-0029.

[16] 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, 2021, doi: 10.1109/ACCESS.2021.3086230.

[17] X. Li *et al.*, “Flowscope: Spotting money laundering based on graphs,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, no. 04, pp. 4731-4738, 2020, doi: 10.1609/aaai.v34i04.5906.

[18] A. Odeh, I. Keshta, and E. Abdelfattah, “Machine Learning Techniques for Detection of Website Phishing: A Review for Promises and Challenges,” *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, NV, USA, 2021, pp. 0813-0818, doi: 10.1109/CCWC51732.2021.9375997.

[19] M. Tiwari, A. Gepp, and K. Kumar, “A review of money laundering literature: the state of research in key areas,” *Pacific Accounting Review*, vol. 32, no. 2, pp. 271-303, 2020, doi: 10.1108/PAR-06-2019-0065.

[20] G. F. Aprilia, “Exploring detection and prevention of money laundering with information technology,” *Journal of Money Laundering Control*, 2023, doi: 10.1108/JMLC-08-2023-0138.

[21] M. Tiwari, J. Ferrill, and D. M. Allan, “Trade-based money laundering: a systematic literature review,” *Journal of Accounting Literature*, 2024, doi: 10.1108/JAL-11-2022-0111.

[22] S. Sterling, “Identifying money laundering: Analyzing suspect financial conduct against the speed, cost, and security of legitimate transactions,” *Journal of Money Laundering Control*, vol. 18, pp. 266-292, 2015, doi: 10.1108/JMLC-08-2014-0025.

[23] S. Dalal, B. Seth, M. Radulescu, C. Secara, and C. Tolea, “Predicting fraud in financial payment services through optimized hyper-parameter-tuned XGBoost model,” *Mathematics*, vol. 10, p. 4679, 2022, doi: 10.3390/math10244679.

[24] M. Mahootiha, A. H. Golpayegani, and B. Sadeghian, “Designing a New Method for Detecting Money Laundering based on Social Network Analysis,” *2021 26th International Computer Conference, Computer Society of Iran (CSICC)*, Tehran, Iran, 2021, pp. 1-7, doi: 10.1109/CSICC52343.2021.9420621.

[25] R. I. T. Jensen and A. Iosifidis, “Fighting Money Laundering With Statistics and Machine Learning,” in *IEEE Access*, vol. 11, pp. 8889-8903, 2023, doi: 10.1109/ACCESS.2023.3239549.

[26] S. N. F. S. M. Nazri, S. H. C.Jailani, S. Zolkaflil, N. H. Ismail, and S. N. S. Yusuf, “The importance of professional skepticism in detecting money laundering: investigating officers’ perspective,” *Asia-Pacific Management Accounting J.*, vol. 18, pp. 275-300, 2023.

[27] N. Pocher, M. Zichichi, F. Merizzi, M. Z. Shafiq, and S. Ferretti, “Detecting anomalous cryptocurrency transactions: An AML/CFT application of machine learning-based forensics,” *Electronic Markets*, vol. 33, p. 37, 2023, doi: 10.1007/s12525-023-00654-3.

[28] W. W. Lo, G. K. Kulatilleke, M. Sarhan, S. Layeghy, and M. Portmann, “Inspection-L: self-supervised GNN node embeddings for money laundering detection in bitcoin,” *Applied Intelligence*, vol. 53, pp. 19406-19417, 2023, doi: 10.1007/s10489-023-04504-9.

[29] D. Dablain, B. Krawczyk and N. V. Chawla, “DeepSMOTE: Fusing Deep Learning and SMOTE for Imbalanced Data,” in *IEEE Transactions on Neural Networks and Learning Systems*, vol. 34, no. 9, pp. 6390-6404, Sept. 2023, doi: 10.1109/TNNLS.2021.3136503.

BIOGRAPHIES OF AUTHORS

Ammar Odeh received his Ph.D. from the University of Bridgeport (UB), USA, in 2015. He is an associate professor at the Department of Computer Science, Faculty of King Hussein School of Computing Sciences, Princess Sumaya University for Technology, Amman, Jordan. His research interests include cybersecurity and cryptography, and the internet of things (IoT). He can be contacted at email: a.odeh@psut.edu.jo.

Anas Abu Taleb is an associate professor in the Department of Computer Science at Princess Sumaya University for Technology, Amman, Jordan. He received a Ph.D. in Computer Science from the University of Bristol, UK 2010, an M.Sc. in Computer Science from the University of the West of England, UK, 2007 and a B.Sc. degree in Computer Science from Princess Sumaya University for Technology, Jordan, 2004. He has published several journal and conference papers on sensor networks. In addition to sensor networks, he is interested in network fault tolerance, routing algorithms, and mobility models. He can be contacted at email: a.abutaleb@psut.edu.jo.