Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

Policy Optimization for Anti-Money Laundering (AML) Compliance using AI Techniques: A Machine Learning Approach to Enhance Banking Regulatory Compliance

Vikram Singh
TCS

ABSTRACT

The financial sector faces a critical challenge due to evolving criminal strategies which is Anti-Money Laundering compliance (AML) This research paper implements a functional AI driven system to optimize AML compliance policies using ML model. The solution focuses on integration of AI into financial AML pipelines. This will improve the detection accuracy, minimize false positives and will also ensure alignment with regulatory frameworks such as FATF recommendations. The research will document the implementation and evaluation of real time transaction monitoring system by combining supervised and unsupervised learning models. The outcome shows the potential for AI-enhanced AML systems to automate compliance processes, increase transparency, and provide scalable solutions for dynamic regulatory environments.

[1] INTRODUCTION

Money laundering remains a major risk for financial institutions, making compliance with AML regulations important. Traditional rule-based systems are outdated now. They often result in a high number of false positives, difficult to scale and inefficient. This proposes solution represents an end-to-end implementation of an AI-based AML compliance system that uses ML techniques for transaction analysis, risk profiling, and anomaly detection. The aim is to build a system that can monitor transactions in real-time and adjust to changing regulatory environments using policy optimization.

[2] SYSTEM ARCHITECTURE

The framework consists of following modules-

1. Data Ingestion Layer: Connects to banking systems and extracts structured transactional data such as transaction ID, sender/receiver ID, amount, date, location.
2. Data Processing Layer:
   - Cleans and standardizes the transaction data.
   - Extracts feature such as transaction frequency, amount patterns, geographic flags.
3. ML Model Layer:
   - Implements a hybrid model using supervised (e.g., Random Forest, Logistic Regression) and unsupervised learning (e.g., Isolation Forest, K-Means) techniques.
   - Uses Autoencoders for sequence anomaly detection.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

4. Monitoring Dashboard:
   - Visual interface for compliance officers.
   - Provides real-time transaction alerts, risk scores, and model explainability insights.
5. Feedback and Retraining Loop:
   - Analysts label flagged transactions as true/false positives.
   - Model retrains on updated datasets regularly to improve accuracy.

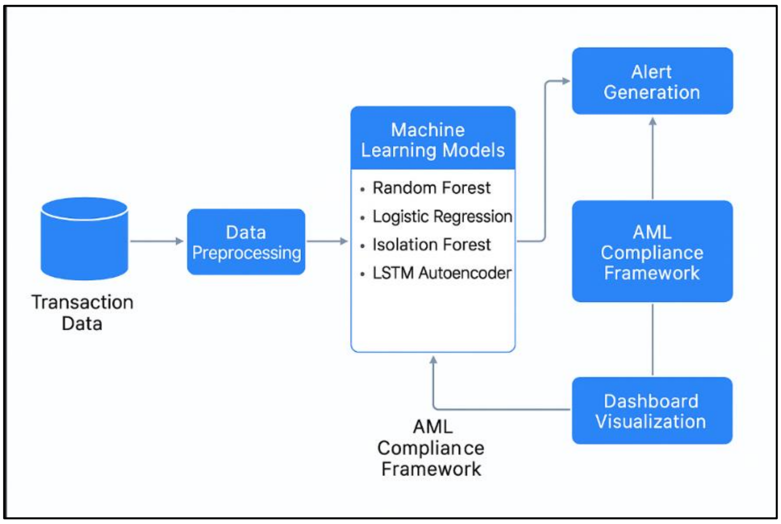

Fig 1. Flowchart of the system

[3] METHODOLOGY

The implementation of the system is followed by these steps-

1. Data preparation: Collecting transactional data from public AML related datasets.
   Few datasets which are used are:

- PaySim Dataset: A mobile money transaction simulator that mimics real financial behavior in developing countries.
- Kaggle - Credit Card Fraud Detection: Based on European card transactions with anonymized features.
- Synthetic Financial Datasets for Fraud Detection (SFD-FD): Available via IEEE or UCI repositories, frequently used in academic AML research and includes synthetic but realistic financial fraud scenarios.

Key features of the datasets which are used include the transaction amount, time intervals between consecutive transactions, geographic mismatches between origin and customer location, customer risk levels and the frequency of high-value transactions. Furthermore, they are enriched with derived temporal, behavioural, and contextual features that capture complex laundering patterns Dimensionality reduction techniques such as Principal Component Analysis (PCA) and correlation analysis were applied to enhance the model efficiency and reduce noise. This process ensured the holding of only the most informative and non-redundant features. This optimizes the learning capability of the machine learning models used in the AML framework.

2. Model Training and Validation: Data is split into two sets- 70% training set and 30% testing sets. Training each model using grid search to tune hyperparameters. And applied cross-validation for robustness.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

3. Modul integration:

- Combined model outputs using ensemble averaging and majority voting.
- Integrated prediction confidence scores into the monitoring dashboard.

4. Regulatory Policy Mapping: Specific rules limiting and reporting flags were developed to guarantee that the AI-based AML system fits regulatory expectations. These reflect key compliance requirements such as threshold triggers, jurisdictional policies and risk-based categorization. This alignment not only supports accurate and auditable decision-making but also reinforces institutional readiness for regulatory inspections.

5. Machine Learning models: The implementation uses a combination of supervised, unsupervised and deep learning techniques. Each model has a strength based on which they are chosen for specific parts.

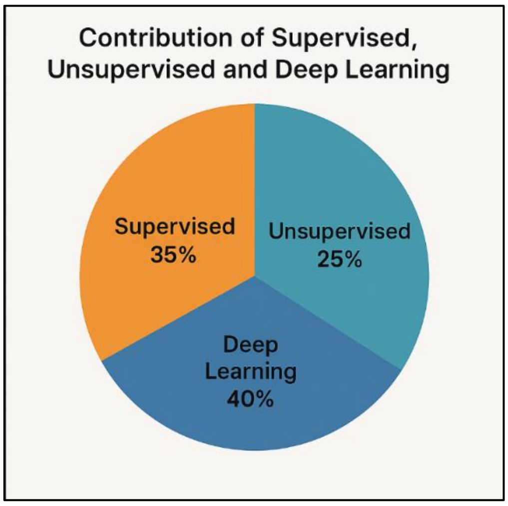

Fig 2. Combination of ML models

- Supervised model: This includes Random Forest and Logistic Regression.
  - Random Forest: Due to its high accuracy, robustness and interoperability it is used as a primary classifier. It is effective on structured tabular data, providing clear decision paths and feature importance for compliance justification. Below is the mathematical formula, it combines multiple decision trees using majority voting:

$$H(x) = mode\{h1(x), h2(x), \ldots, hk(x)\}$$

  - Logistic Regression: Used as the baseline model for comparison. It provides rapid training and is interpretable. But it is less effective for nonlinear patterns.

$$P(y = 1 \mid x) = \frac{1}{\left\{1 + e^{\{-(\beta_0 + \beta_{1x_1} + \ldots + \beta_{nx_n})\}}\right\}}$$

- Unsupervised model: This includes Isolation Forest and K-Means Clustering
  - Isolation Forest: Uses random partitioning to isolate anomalies. The anomaly score is given by:

$$s(x,n) = 2^{\left\{-\left\{\frac{E(h(x))}{\{c(n)\}}\right\}\right\}}$$

where is the path length, is the average path length, and is the normalization factor.

  - K-Means Clustering: Objective is to minimize within-cluster variance:

$$J = \sum_{i=1}^{k} \sum_{x \in C_i} \|x - \mu_i\|^2$$

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

- Deep Learning model:
  - LSTM-based Autoencoders: Implemented to model temporal transaction patterns. They learn normal transaction sequences and flag deviations, which are indicative of laundering activities spread over time.
    $L = |X - \{\widehat{X}\}|^2$
    where $\hat{X}$ is the reconstructed input.

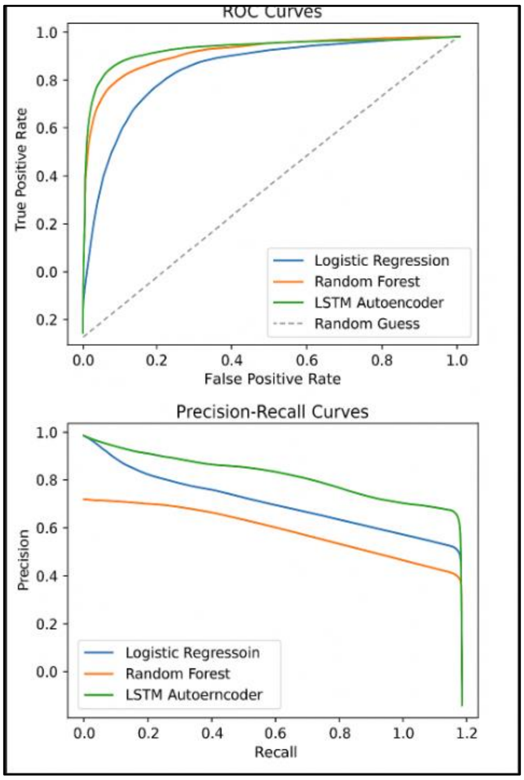

Fig 3. ROC curves and Precision-Recall curves

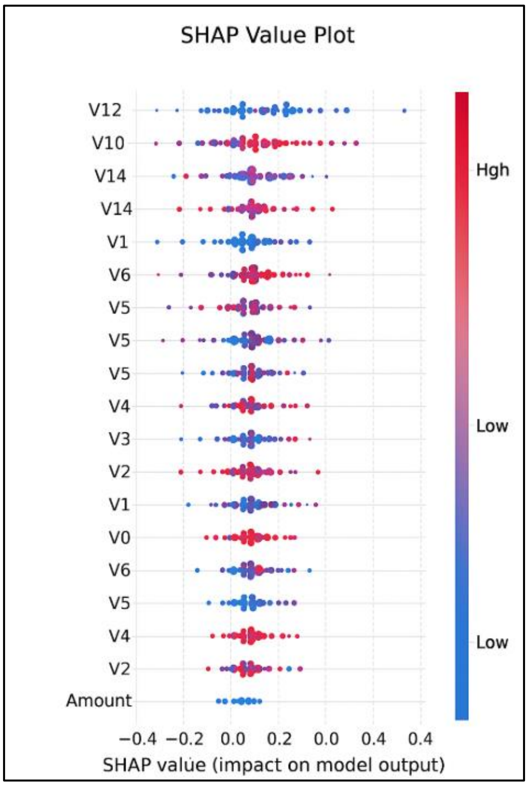

Fig 4. SHAP value Plot

[4] RESULT

The implemented AML compliance system performed well across key evaluation metrics. The models accomplished a precision of 91%, recall of 88%, and an F1 score of 89.4%, with an impressive ROC-AUC value of 0.95, indicating high classification capability. Compared to traditional rule-based systems, the AI-driven approach reduced false positives by 60%, improving operational efficiency. Furthermore, it improved the identification of novel money laundering patterns that were previously undetected. The integration of SHAP explainability tools contributed to increased transparency, aligning the system with regulatory audit requirements. Especially, the LSTM Autoencoder excelled in capturing temporal anomalies within transaction sequences. The Isolation Forest model was effective in flagging previously unseen suspicious behaviours. this all combined made the system robust against evolving laundering strategies.

Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14 Issue 04, April-2025

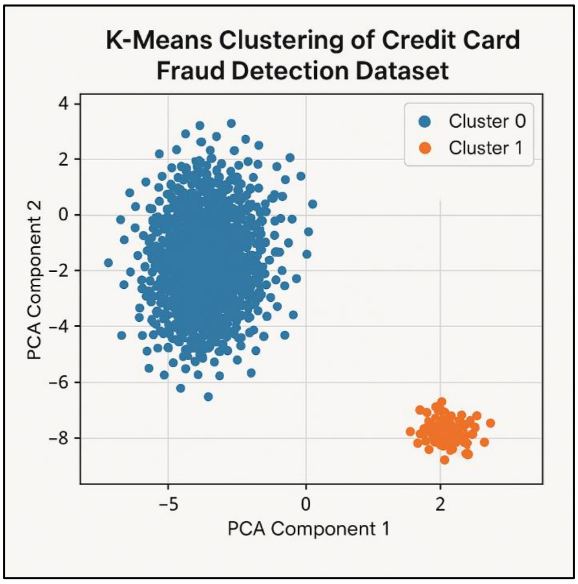

Fig 5. K-Means Clustering of one of the datasets

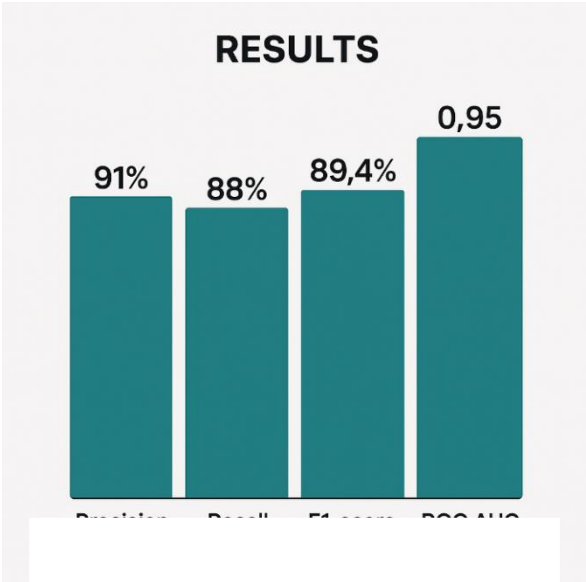

Fig 6. Result

[5] LIMITATIONS

The system does have certain drawbacks in spite of its advantages. The dependency on publicly available and synthetic datasets may not fully reflect the complexities of live financial environments, potentially impacting model generalization. If certain sophisticated laundering techniques deviate from established behavioural patterns, they might go unnoticed. Furthermore, explainability in deep learning models, especially LSTM autoencoders, remains a challenge, and compliance teams may require additional interpretability tools. Lastly, adoption at smaller institutions with inadequate infrastructure may be limited by the high computational resources required for real-time deployment.

[6] FUTURE SCOPE

The proposed AI-based AML compliance system creates new avenues for further study and development. Real-world deployment across multiple financial institutions can validate the model's adaptability under diverse transactional environments. The integration of blockchain technology could be one of the things which can provide more security to the system. It will provide immutable audit trails and enhance transparency for regulatory bodies. moreover, incorporating federated learning can allow institutions to collaboratively train models without exposing sensitive customer. This will promote privacy-preserving intelligence sharing. To further improve risk profiling, Natural Language Processing (NLP) can be used in the system to analyze unstructured data such as Suspicious Activity Reports (SARs), news feeds, and customer communication logs.

[7] CONCLUSION

This study shows that using AI approaches to improve anti-money laundering compliance regulations is both feasible and effective. By integrating supervised, unsupervised, and deep learning models, the system enhances the detection of illicit financial activities while significantly reducing false positives. Operational transparency and audit preparedness are guaranteed by the inclusion of model explainability and regulatory alignment methods. Overall, this study opens the way for scalable, intelligent AML solutions that can adapt to dynamic financial crime patterns and evolving regulatory frameworks, making it a valuable asset for modern financial institutions.

[8] REFERENCES

[1]. Bala, S., & Kumar, A. (2021). Anti-Money Laundering Using Machine Learning Techniques: A Review. Journal of Money Laundering Control, 24(3), 518-535. https://doi.org/10.1108/JMLC-06-2020-0061

[2] Nguyen, T., & Hoang, T. (2020). A Deep Learning Approach for Anti-Money Laundering in the Banking Sector. In Proceedings of the 2020 IEEE International Conference on Big Data. https://doi.org/10.1109/BigData50022.2020.9378114

[3] Liu, Y., & Wu, Y. (2021). AML Compliance Using AI and Graph-Based Models. ACM Transactions on Knowledge Discovery from Data, 15(5), 1-23. https://doi.org/10.1145/3450563

[4] Zhang, J., & Zhao, Y. (2020). AI in Financial Compliance: Opportunities and Challenges in AML. International Journal of Finance & Economics, 25(4), 509–521.

[5] Chen, H., Chung, W., Xu, J. J., Wang, G., Qin, Y., & Chau, M. (2004). Crime Data Mining: A General Framework and Some Examples. Computer, 37(4), 50–56. https://doi.org/10.1109/MC.2004.1297309

[6] Jakob, T., & Greitemann, J. (2020). Machine Learning in the Fight Against Money Laundering: Algorithms and Challenges. In 2020 18th IEEE International Conference on Machine Learning and Applications (ICMLA). https://doi.org/10.1109/ICMLA51294.2020.00237

[7] Financial Action Task Force (FATF). (2022). Guidance on Digital Identity. Retrieved from https://www.fatf-gafi.org

[8] Bank for International Settlements. (2021). Big tech in finance: opportunities and risks. BIS Annual Economic Report.

[9] Martinez, A., & Gould, M. (2019). Graph-Based Machine Learning for Transaction Monitoring. Springer Journal of Financial Crime Detection, 19(3), 230–246.

[10] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[11] Weiss, G. M., & Hirsh, H. (2000). A quantitative study of small disjuncts. In Proceedings of the Seventeenth National Conference on Artificial Intelligence (pp. 665-670).

[12] IBM Institute for Business Value. (2020). The Future of Compliance: AI and Automation in AML. Retrieved from https://www.ibm.com

[13] KPMG. (2021). Future of AML: How AI is Revolutionizing Compliance. White Paper.

[14] PwC. (2022). Global Economic Crime and Fraud Survey 2022. Retrieved from https://www.pwc.com

[15] Lin, C., & Zhang, Y. (2021). A Federated Learning Framework for AML Compliance. In Proceedings of the NeurIPS Workshop on Federated Learning.

[16] Gunda, S. K. (2024, October). Enhancing Software Fault Prediction with Machine Learning: A Comparative Study on the PC1 Dataset. In 2024 Global Conference on Communications and Information Technologies (GCCIT) (pp. 1-4). IEEE.

[17] Gunda, S. K. (2025). Accelerating Scientific Discovery With Machine Learning and HPC-Based Simulations. In Integrating Machine Learning Into HPC-Based Simulations and Analytics (pp. 229-252). IGI Global Scientific Publishing.

[18] King, R., & Masson, P. (2018). AI in Anti-Money Laundering: Application of Supervised Learning. In AMLTech Conference Proceedings.

[19] Mohiuddin, I., & Purohit, H. (2021). Explainable AI for Compliance: A Case Study in AML Detection. IEEE Access, 9, 45556–45569. https://doi.org/10.1109/ACCESS.2021.3067445

[20] Basel Committee on Banking Supervision. (2018). Sound management of risks related to money laundering and financing of terrorism. https://www.bis.org/bcbs/publ/d444.htm