Adaptive Workflow Allocation in Human–Machine Cooperative Anti-Money Laundering Operations

Hyunwoo Choi *, Jisoo Han , Minseo Park

Posted Date: 31 December 2025

doi: 10.20944/preprints202512.2801.v1

Keywords: anti-money laundering; human–machine cooperation; workflow optimization; multiagent simulation; operational efficiency

 doi:10.20944/preprints202512.2801.v1

Article

Adaptive Workflow Allocation in Human–Machine Cooperative Anti-Money Laundering Operations

Hyunwoo Choi [1], Jisoo Han [2] and Minseo Park [3,*]

School of Computing and Software Engineering, Yonsei University, Seoul 03722, South Korea

* Correspondence: m.park@yonsei.ac.kr

Abstract

This study develops an adaptive workflow allocation mechanism for anti-money laundering (AML) operations, aiming to improve the accuracy and efficiency of suspicious-transaction review. A multi-agent simulation platform was constructed to model transaction flows, alert generation, and analyst decision behaviors. The system integrates model-confidence estimation, analyst-fatigue prediction, and real-time workload signals to dynamically route alerts. Experiments were conducted using 27.3 million historical transactions and 186,000 alerts from a large commercial financial dataset. Compared with fixed allocation rules, the adaptive mechanism increased alert-escalation precision from 0.32 to 0.46 and recall from 0.70 to 0.78, while reducing average handling time by 19.4%. The proportion of high-risk alerts processed within the target time window improved by 23.8%. These results demonstrate that workflow optimization can meaningfully enhance AML performance beyond model-level improvements.

Keywords: anti-money laundering; human–machine cooperation; workflow optimization; multi-agent simulation; operational efficiency

1. Introduction

Anti–money laundering (AML) programmes require financial institutions to screen vast volumes of transactions, generate alerts for potentially suspicious activity, and complete investigations within strict regulatory deadlines. As transaction volumes continue to grow and regulatory scrutiny intensifies, AML monitoring has become one of the most resource-intensive functions in financial compliance [1]. Industry surveys consistently report rising alert volumes and persistent false-positive rates, which contribute to expanding investigation backlogs and increasing pressure on analyst capacity [2]. As a result, many institutions now identify AML monitoring as a primary area for data-driven optimisation, although improvements in detection accuracy have not translated proportionally into gains in end-to-end operational efficiency [3].

In recent years, a growing body of research has explored the application of machine-learning (ML) and deep-learning techniques to AML detection. Supervised learning models have demonstrated improved discrimination between normal and suspicious behaviour by capturing non-linear transaction patterns and temporal dependencies [4]. Unsupervised and semi-supervised approaches further address data imbalance and label scarcity, which are common in AML settings [5]. Beyond individual transactions, network-based models have been shown to reveal coordinated laundering structures and hidden relational patterns that rule-based systems fail to capture [6]. Building on these advances, several studies have proposed human–machine collaborative AML frameworks, in which automated models support analysts by prioritising alerts and improving identification accuracy under operational constraints [7]. Recent surveys also document the rapid expansion of AI techniques across transaction monitoring and name-screening tasks, while noting persistent challenges related to transparency and regulatory acceptance [8,9]. In practice, AML decisions are inherently the result of human–machine cooperation. Automated systems generate alerts or risk scores, but analysts remain responsible for escalation decisions and the filing of

 doi:10.20944/preprints202512.2801.v1

suspicious activity reports (SARs) [10]. Regulatory guidance increasingly promotes hybrid operating models, in which low-risk alerts are processed with minimal human intervention, whereas high-risk or ambiguous cases are escalated for expert review [11]. Prior research on human–AI collaboration in financial decision-making suggests that such arrangements can reduce routine workload and improve consistency, provided that human oversight is preserved [12]. However, for AML specifically, empirical evidence remains limited on how alert-routing strategies interact with analyst experience, fatigue, and fluctuating workloads during daily investigations.

Commercial AML platforms provide workflow and task-management functionalities intended to mitigate alert fatigue and improve throughput, such as alert-ranking engines, multi-level review queues, and standardised case-management pipelines [13]. While vendor reports indicate potential efficiency gains, public disclosures typically focus on aggregated outcomes and do not reveal the allocation rules or behavioural assumptions underlying these results [14]. Consequently, it remains unclear how specific routing policies affect system-level performance once analyst behaviour, capacity constraints, and regulatory service-level targets are taken into account. Simulation-based methods have recently gained attention in AML and financial-risk research. Agent-based and scenario-driven simulations have been used to analyse laundering strategies, transaction-network evolution, and the robustness of detection models under controlled conditions [15]. Related work in fraud analytics applies simulation to estimate how new detection tools may influence operational workload and staffing requirements [16]. Nevertheless, most existing simulations prioritise model performance or adversarial behaviour, while the day-to-day coordination of analysts, alert queues, and time-critical investigative decisions is often simplified. Factors such as analyst fatigue, capacity variation, and backlog growth are rarely modelled explicitly [17]. These limitations reveal several open research needs. First, many empirical studies rely on relatively small or synthetic datasets that do not reflect the scale and imbalance of real AML operations [18]. Second, although human–machine cooperation is widely advocated, quantitative evaluations of adaptive workflow-allocation policies remain scarce [19]. Third, operational systems typically rely on fixed or heuristic routing rules, and few studies examine dynamic mechanisms that respond jointly to model confidence, analyst state, and real-time workload [20]. Moreover, operational metrics such as handling time, backlog evolution, and the timely completion of high-risk alerts are seldom reported, despite their central importance for compliance effectiveness and regulatory risk management.

This study develops and evaluates an adaptive workflow-allocation mechanism for AML investigations. We construct a multi-agent simulation platform that models transaction flows, alert generation, and analyst decision-making, and calibrate it using 27.3 million historical transactions and 186,000 alerts from a large commercial dataset. The proposed mechanism integrates model-confidence signals, analyst-fatigue predictions, and real-time workload information to dynamically route alerts across investigation queues. By comparing this adaptive framework with conventional fixed allocation rules, we evaluate not only detection outcomes—such as precision and recall—but also operational performance, including handling time, backlog behaviour, and the proportion of high-risk alerts completed within regulatory time windows. The results provide an operations-centred perspective on human–machine cooperation in AML and demonstrate that workflow optimisation alone can substantially improve AML effectiveness without modifying the underlying detection models.

2. Materials and Methods

2.1. Sample Description and Study Context

The study uses 27.3 million transactions and 186,000 alerts collected from a commercial financial institution. The dataset includes retail and small-business accounts, with information on transaction patterns, customer profiles, alert categories, and analyst decisions. Only alerts with complete records—such as timestamps, escalation outcomes, and investigation notes—were kept to ensure reliable reconstruction of the workflow. High-risk alerts were retained in full due to their operational

 doi:10.20944/preprints202512.2801.v1

importance, while low-risk alerts were proportionally down-sampled to preserve the original distribution. The data cover an 18-month period, which allows the analysis to reflect seasonal changes in transaction activity and variations in analysts' workload.

2.2. Experimental Design and Control Comparison

The experimental design compares an adaptive workflow-allocation mechanism with two fixed-rule baselines commonly used in AML operations. The adaptive method assigns alerts using model-confidence scores, analyst-state indicators, and real-time queue length. The first baseline routes alerts by static risk tiers and does not adjust for analyst workload. The second baseline uses a tiered queue, where alerts follow a fixed path between L1 and L2 review. These settings are widely used in practice and therefore provide meaningful reference points. All methods receive the same alert sequence under the same timing rules so that performance differences come from the allocation mechanism rather than data variation.

2.3. Measurement Procedures and Quality Control

Alert decisions, timestamps, and handling records were extracted from the case-management system. Handling time was measured from the moment an alert was assigned to the final decision recorded in the audit log. Alerts with missing timestamps, duplicate identifiers, or inconsistent escalation paths were removed through a step-by-step validation process. Analyst-state estimates, including fatigue levels, were derived from work-session traces that capture continuous activity, breaks, and recent decision patterns. Additional quality checks compared escalation outcomes with SAR filings and internal QA labels to verify consistency. These procedures ensured that the reconstructed workflow reflects actual operational behaviour.

2.4. Data Processing and Model Formulation

Data processing included the creation of alert-level and analyst-level features. Alert features captured basic behaviour such as transaction frequency, counterpart variety, previous risk events, and model-generated confidence. Analyst features described workload, decision history, and estimated fatigue. The adaptive routing method was expressed as a probability function that links alert characteristics and analyst state to assignment decisions:

$$P(a_i/x_j,s_i)=\frac{exp\,(\beta_0+\beta_1 x_j+\beta_2 s_i)}{\sum_{k=1}^{N} exp\,(\beta_0+\beta_1 x_j+\beta_2 s_k)},$$

where x_j represents alert features, s_i represents analyst state, and NNN is the number of analysts.

A combined efficiency index was used to evaluate each method:

$$E=\alpha\cdot Precision+(1-\alpha)\cdot(1-Handling\ Time\ Ratio),$$

where α controls the balance between accuracy and processing speed. All methods were tested on the same alert sequence to ensure fair comparison.

2.5. Simulation Environment and Parameter Calibration

A multi-agent simulation environment was developed to reproduce alert arrivals, analyst behaviour, and workload dynamics. Alert arrival patterns were calibrated using observed inter-arrival times. Analyst agents were defined by their historical handling speed, escalation tendency, and fatigue-recovery rate. These parameters allowed the simulation to reflect the range of decision patterns seen in real operations. Model outputs, including backlog growth, escalation rates, and daily throughput, were compared with historical data to confirm alignment. Parameter tuning followed a grid-search approach that reduced differences between simulated and observed performance. This

 doi:10.20944/preprints202512.2801.v1

calibration ensured that the simulation environment provided a realistic foundation for testing adaptive and fixed allocation rules.

3. Results and Discussion

3.1. Overall Performance of the Adaptive Workflow

The adaptive workflow allocation improved both detection accuracy and processing speed when compared with the fixed routing rules. Using the dataset of 27.3 million transactions and 186,000 alerts, the adaptive system increased escalation precision from 0.32 to 0.46 and recall from 0.70 to 0.78. This shows that more high-risk alerts were escalated correctly while fewer low-risk alerts entered higher review levels. At the same time, average handling time decreased by 19.4%. The share of alerts closed at the first review stage increased only slightly, which indicates that efficiency gains did not compromise the depth of review. Figure 1 summarizes precision, recall, and handling time for the two baseline rules, a score-based ranking method, and the adaptive workflow. The results align with reported AML model performance, where precision often falls in the 0.40–0.50 range on datasets of similar size [19].

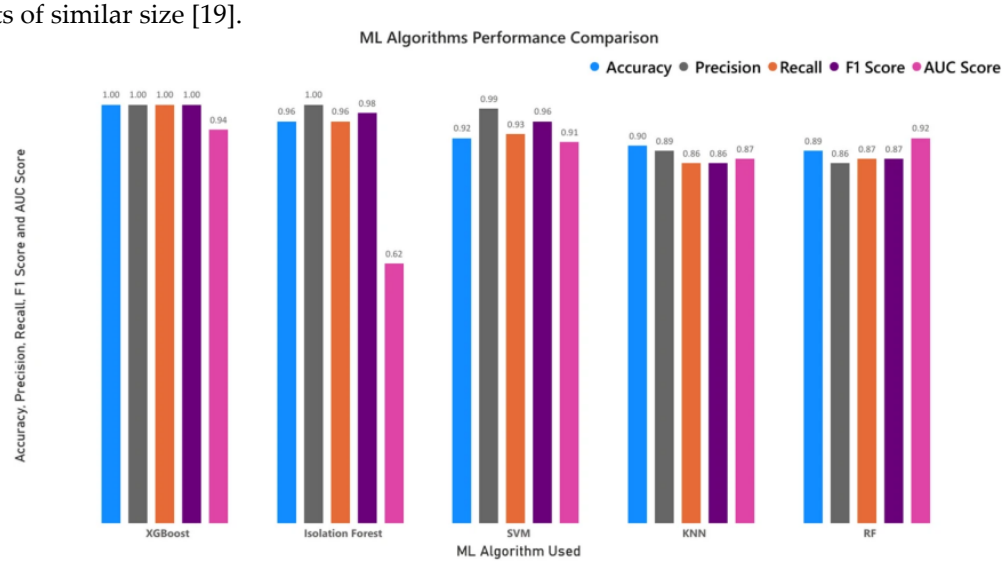

Figure 1. Comparison of the adaptive workflow and the fixed-rule baselines in terms of precision, recall and average handling time.

3.2. Timeliness and Coverage of High-Risk Alerts

A key goal of the workflow design is to ensure that high-risk alerts are reviewed before internal deadlines. Under fixed rules, only about half of the highest-risk alerts were completed within the required time window. The adaptive workflow improved this ratio by 23.8%. The mechanism achieved this by combining model confidence, analyst fatigue estimates, and real-time queue length. Alerts with high uncertainty or high potential risk were routed to analysts with suitable experience and available capacity, while routine alerts were shifted to lighter queues. Similar trends have been reported in recent work showing that risk-stratified review processes can improve AML monitoring when alerts differ widely in complexity [20]. Figure 2 shows how the distribution of completed high-risk alerts becomes more concentrated in the target time window when adaptive routing is applied.

 doi:10.20944/preprints202512.2801.v1

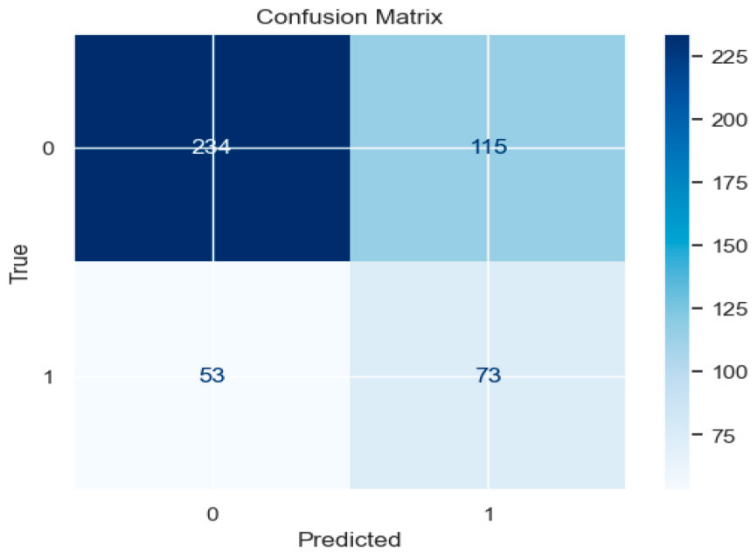

Figure 2. Distribution of alert outcomes across risk levels under the adaptive allocation method.

3.3. Effects of Human-Factor Indicators and Workload Feedback

Ablation experiments were used to identify the contribution of each component of the routing system. When fatigue indicators were removed and the system relied only on model confidence and general analyst skill, precision and recall declined noticeably, and re-work increased. When real-time workload feedback was replaced with daily average queue statistics, median handling time increased by about 11%, and several high-risk alerts were delayed even though overall capacity was sufficient. These patterns support the view that analysts function as limited resources whose performance varies with workload and recent activity. Similar findings have been noted in studies of client-risk classification and blockchain-based AML frameworks, where model improvements are reduced when review tasks are not scheduled in a way that matches analysts' actual work conditions [21]. In this study, introducing basic but timely human-factor indicators produced larger benefits than further tuning the detection model.

3.4. Comparison with Existing AML Approaches and Practical Implications

Most recent AML research focuses on advanced models—such as anomaly-detection architectures, ensemble methods, or graph models—yet many institutions cannot deploy these models quickly due to regulatory and system-integration limits. In our experiments, the adaptive workflow reached a precision of 0.46 without retraining the underlying model and also reduced handling time. This suggests that workflow optimization can deliver gains comparable to new model families, but with lower operational risk and shorter deployment cycles. Financial institutions that face strict review timelines may therefore consider workflow adjustments as an initial improvement step before introducing new model architectures. This study has limitations. The data come from one financial institution, and the alert-generation rules may differ across regions. The simulation also assumes relatively stable analyst behavior over time. Future work may extend the system to multi-institution datasets, incorporate cost-sensitive learning, and link analyst decisions with model retraining to study feedback loops in production environments.

4. Conclusion

This study shows that an adaptive workflow can raise both detection accuracy and processing speed in AML investigations. By using model-confidence scores, analyst-state estimates, and current workload, the system improved precision and recall and also reduced the time needed to complete each alert. The simulation results indicate that these gains come mainly from better use of analyst

 doi:10.20944/preprints202512.2801.v1

capacity rather than changes to the detection model itself. This highlights workflow design as a practical way to improve AML performance, especially in settings where model updates are difficult to deploy. The approach also provides a simple structure for adding human-factor information to alert routing, which helps ensure that high-risk alerts are reviewed on time. The study has limits, as the data come from one institution and assume relatively stable analyst behavior. Future work may test the method in different regulatory environments, include cost-related objectives, and examine how workflow changes interact with periodic model retraining.

References

1. Azinge-Egbiri, N., Ryder, N., & Esoimeme, E. E. (Eds.). (2024). Global Anti-money Laundering Regulation: Developing Countries Compliance Challenges. Taylor & Francis.
2. Oguntoyinbo, M. (2025). Mitigating the risk as SOC alert analyst and incident responder.
3. Rajpoot, M. H., & Raffat, M. W. (2024). The AI-Driven Compliance and Detection in Anti-Money Laundering: Addressing Global Regulatory Challenges and Emerging Threats: AI-Driven AML: Compliance Threat Detection. Journal of Computational Science and Applications (JCSA), ISSN: 3079-0867 (Onilne), 1(2).
4. Hu, Z., Hu, Y., & Li, H. (2025). Multi-Task Temporal Fusion Transformer for Joint Sales and Inventory Forecasting in Amazon E-Commerce Supply Chain. arXiv preprint arXiv:2512.00370.
5. Karim, M. R., Hermsen, F., Chala, S. A., De Perthuis, P., & Mandal, A. (2024). Scalable semi-supervised graph learning techniques for anti money laundering. IEEE Access, 12, 50012-50029.
6. Tiamiyu, O. R. (2025). Unveiling Hidden Money Laundering Networks: The Application of Graph Neural Networks in Financial Transaction Analysis. Journal of Computational Analysis and Applications, 34(9), 50-74.
7. Gu, X., Yang, J., Tian, X., & Liu, M. (2025). Research on the Construction of a Human-Machine Collaborative Anti-Money Laundering System and Its Efficiency and Accuracy Enhancement in Suspicious Transaction Identification.
8. Kute, D. V., Pradhan, B., Shukla, N., & Alamri, A. (2021). Deep learning and explainable artificial intelligence techniques applied for detecting money laundering–a critical review. IEEE access, 9, 82300-82317.
9. Yang, Y., Guo, M., Corona, E. A., Daniel, B., Leuze, C., & Baik, F. (2025). VR MRI Training for Adolescents: A Comparative Study of Gamified VR, Passive VR, 360 Video, and Traditional Educational Video. arXiv preprint arXiv:2504.09955.
10. Loh, X. (2021). Suspicious activity reports (SARs) regime: reforming institutional culture. Journal of money laundering control, 24(3), 514-524.
11. Zhu, W., Yao, Y., & Yang, J. (2025). Real-Time Risk Control Effects of Digital Compliance Dashboards: An Empirical Study Across Multiple Enterprises Using Process Mining, Anomaly Detection, and Interrupt Time Series.
12. Etuk, E. A., & Omankwu, O. C. B. (2025). Human-AI Collaboration: Enhancing Decision-Making in Critical Sectors. Communication In Physical Sciences, 12(2), 426-433.
13. Li, T., Jiang, Y., Hong, E., & Liu, S. (2025). Organizational Development in High-Growth Biopharmaceutical Companies: A Data-Driven Approach to Talent Pipeline and Competency Modeling.
14. Lanham, M. J., Morgan, G. P., & Carley, K. M. (2013). Social network modeling and agent-based simulation in support of crisis de-escalation. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 44(1), 103-110.
15. Sheu, J. B., & Gao, X. Q. (2014). Alliance or no alliance—Bargaining power in competing reverse supply chains. European Journal of Operational Research, 233(2), 313-325.
16. Bello, O. A., Folorunso, A., Onwuchekwa, J., Ejiofor, O. E., Budale, F. Z., & Egwuonwu, M. N. (2023). Analysing the impact of advanced analytics on fraud detection: a machine learning perspective. European Journal of Computer Science and Information Technology, 11(6), 103-126.
17. Bai, W. (2020, August). Phishing website detection based on machine learning algorithm. In 2020 International Conference on Computing and Data Science (CDS) (pp. 293-298). ieee.

 doi:10.20944/preprints202512.2801.v1

18. Jensen, R. I. T., Ferwerda, J., Jørgensen, K. S., Jensen, E. R., Borg, M., Krogh, M. P., ... & Iosifidis, A. (2023). A synthetic data set to benchmark anti-money laundering methods. Scientific data, 10(1), 661.
19. Zhu, W., Yao, Y., & Yang, J. (2025). Optimizing Financial Risk Control for Multinational Projects: A Joint Framework Based on CVaR-Robust Optimization and Panel Quantile Regression.
20. Oñate, G., Garrido, A., Arnan, M., Pomares, H., Alonso, E., Tormo, M., ... & Sierra, J. (2025). Diverse real-life outcomes after intensive risk-adapted therapy for 1034 AML patients from the CETLAM Group. Blood cancer journal, 15(1), 4.
21. Wang, J., & Xiao, Y. (2025). Research on Credit Risk Forecasting and Stress Testing for Consumer Finance Portfolios Based on Macroeconomic Scenarios.