Sarcouncil Journal of Applied Sciences

**ISSN(Online): 2945-3437**

Volume- 06| Issue- 04| 2026

**Research Article** **Received:** 05-03-2026 | **Accepted:** 30-03-2026 | **Published:** 16-04-2026

# AI-Driven Risk Profiling and Compliance Remediation System Based on Behavioural Pattern Analysis

***Sanjay Chandrakant Vichare***

*N.L. Dalmia Institute of Management Studies and Research, Mumbai, Maharashtra, India*

**Abstract:** Risk profiling using AI has become a topical area in terms of compliance where the volume of transactions, the need to minimize fraud, identity abuse, and increased regulatory pressures are growing. The impossibility to detect the emergent misconduct, hidden abnormalities, and context-specific violation of the compliance rules by fixed rules and regular audits has made the analysis of the behavioural patterns particularly topical. The review forms a conceptual framework of how behavioural risk detection is related to explainable and proportionate compliance remediation. It covers the significant techniques like anomaly detection, sequence learning, graph-based modelling, cost-sensitive classification, and behavioural biometrics and also refers to the problem of explainability, fairness, privacy, and deployment. The literature review has shown that behavioural cues will be effective to increase sensitivity to the emergent threats, especially when dynamics and relational trends are considered. Nevertheless, an imbalanced performance persists in the case of class imbalance, concept drift, delayed labels and institution specific operating conditions. Low interpretability weakens governance structures, low benchmark comparability and weak integration of risk scoring and downstream remediation are some of the weaknesses that have been reported. The advancement in this regard does not simply involve precision in risk identification, but also systems that have the ability to relate the evidence of risks to auditable and visible compliance action that is reasonable.

**Keywords:** Anomaly detection; behavioural analytics; compliance remediation; explainable artificial intelligence; risk profiling.

## INTRODUCTION

Risk profiling has become a critical component of compliance operations in the banking, payment, insurance, and digital platforms to the ranks of highly regulated industries. The previous compliance monitoring systems were relatively grounded on predetermined thresholds, business rules which needed to be manually created and fixed reviewing cycles. This method brought procedural stability. But it rarely worked when behaviour changed in an adversarial way, or when the appearance of suspicious behaviour was only noticed with time, or when seemingly innocent events turned out to be important only with regard to each other. The point that the traditional controls no longer have a high discriminatory value under high volume, heterogeneous and rapidly changing data has been proved many times in the broader literature on fraud and risk [Ngai, E. W. *et al*., 2011]. Statistical fraud detection studies have also demonstrated that abnormal behaviour is hardly manifested in terms of singular definitive occurrence, in most instances it is manifested in terms of frequency, sequence, timing, amount, destination or in the form of relationship structure violations [Bolton, R. J., & Hand, D. J. 2002]. This change toward single-event inspection to behavioural interpretation provided a fertile platform to artificial intelligence techniques, especially machine-learning systems that are able to capture nonlinear dependence and subtle anomaly patterns.

Behavioural pattern analysis broadens the scope of risk assessment. This type of analysis concentrates on patterns of behaviour over time, channel and interaction networks as opposed to concentrating on customer attributes or existing categories of risk. The history of transactions, device history, login patterns, location changes, session patterns, typing patterns, mouse movement, merchant relationship links and customer relationships can all act as behavioural indicators. In the financial sector, such information can support earlier detection of account takeover, mule activity, collusion, layering behaviour, and payment fraud. Similar indicators aid in insider threat detection, policy violation detection and continuous authentication in the context of cybersecurity-related compliance environments. It has been pointed out by the anomaly detection literature that the value of such signals lies not merely in increased data volume, but in greater temporal and relational granularity which reveals changing patterns of misconduct unreachable to more traditional profiling methods [Ahmed, M. *et al.*, 2016]. With the further digitization of regulated sectors, behavioural data increasingly serves as an interface between compliance intelligence and operational events.

It has also transformed the nature of research since risk profiling can no longer be evaluated solely on predictive performance. A model that raises red





***Corresponding Author:** Sanjay Chandrakant Vichare

DOI- https://doi.org/10.5281/zenodo.19613204

flags on suspicious behaviour and provides no plausible justification and a channel to remedial action rather than a benefit of operation may become a governance burden rather than an operational benefit. The explainable AI scholarship has found a conflict between the complex predictive designs and accountability demands of high-stakes decision systems [Adadi, A., & Berrada, M. *et al*., 2018]. The heightening of alert, the blocking of transactions, the outreach to the customers or the closing of a case may often have to be justified to the internal control authorities and external regulators by a compliance staff. Risk scores that are unrelated and cannot be linked to explicit feature importance, time causality and institutional policy are therefore harmful to adoption. What makes the problem even more complicated is that behavioural pattern analysis is dependent on high-dimensional data streams whose latent representations do not readily have a translation to compliance narratives. Design of interpretability, auditability and human review has, in turn, become the subject of scholarly discussion as opposed to implementation issues of secondary concern.

The recent literature on fintech and risk management also suggests that explainable AI is particularly crucial when machine-learning outputs trigger business restrictions, customer friction, or the distribution of resources to investigations [Bussmann, N. *et al*., 2020]. Risk profiling systems are not used alone, with each score affecting case queues, remediation processes, screening intensity, identity re-verification, or account blocking. In practice, a compliance remediation system should be able to transform the detected anomalies into commensurate and justifiable action. This may include gentle reminders about document update, the dynamic adjustment of transaction limits, increased monitoring limits, the creation of analyst cases, recalibration of rules, or retraining of models depending on the results of the investigators. However, the scholarship remains skewed toward behaviour identification rather than remediation design. A lot of research maximizes the classification accuracy and little effort is paid to prioritization of alerts, calibration of remediation thresholds or corrective actions based on subsequent model outputs or investigator feedback. This is a gap that is consequential since a weak remediation logic may increase the cost of operation, loss of customers and fairness issues despite the detector being seemingly correct.

Methodological pluralism is another significant characteristic of the present-day picture. The sequence models, graph-based models, anomaly detectors and cost-sensitive classifiers are used to model order and time dependencies of events, relationships between entities and transactions, detect deviations of expected behaviour and deal with serious class imbalance typical of fraud and compliance data. Behavioural biometrics provides a unique source of identity-related data by evaluating the current interaction trends in comparison to a legitimate user profile. Coexisting with this diversification, researchers have reported a sequence of repetitive limitations: scanty labels, delayed fraud confirmation, non-stationary behaviour, institution-specific feature spaces, privacy constraints, and asymmetric error costs [Ngai, E. W. *et al*., 2011; Ahmed, M. *et al*., 2016]. These limitations make cross-study comparison difficult. A high-recall model may impose an unacceptable false-positive burden in live compliance environments; a highly accurate model can fail to detect new patterns of attack; a seemingly robust system can fail rapidly once the opponents evolve. Significant review thus cannot be done without more than cataloguing algorithms. It demands a close consideration of the assumptions about operations, regimes of data, measures of evaluation, and institutional setting.

This topic is broadly applicable across financial fraud analytics. Behavioural risk profiling is an interdisciplinary field of computer science, information systems, cybersecurity, regulatory research, digital identity management, ethics, and operations research, which is driven by AI. It also raises some significant legal and managerial issues of proportionality, due process, data minimization and model governance. Behavioural intelligence can enhance institutional resilience and improve the detection of concealed malpractices, but similar mechanisms can lead to transparency, excessive policing, and discriminatory action in the presence of weak governance structures. This field therefore has a distinctly interdisciplinary character where both technical capability and compliance legitimacy should be constructed simultaneously as opposed to in sequence.

This review aims to analyse the conceptualization, construction, and assessment of AI-based risk profiling and compliance remediation systems based on behavioural pattern analysis reported in prior studies. In addition, this paper proposes a structured framework to link risk detection with remediation actions. The discussion begins with a



**Publisher: SARC Publisher**

review of the main strands of the literature and then develops the conceptual and methodological approaches. A subsequent section analyses reported findings and their implications for operational compliance. The review then outlines future research directions and concludes with the major implications for scholarship and practice.

## LITERATURE REVIEW

The literature on AI-based risk profiling has evolved in a number of more or less overlapping directions: financial fraud detection, anomaly detection, behavioural biometrics, online banking security and explainable risk analytics. Foundational review studies provided the intellectual foundation, demonstrating that financial fraud is a high-dimensional classification problem, with rare positive instances, as well as concept drift, delayed labels, and heterogeneous feature spaces [West, J., & Bhattacharya, M. 2016], [Abdallah, A. *et al*., 2016]. In that context, researchers gradually moved away from customer-level profiling to behaviour-based modelling that captures behaviour over time. This change can be observed in sequence classification studies, transactional pattern mining and real-time fraud scoring case studies. Studies such as [Jurgovsky, J. *et al*., 2018] and [Carcillo, F. *et al*., 2021] show that suspicious behaviour often appears as temporal anomalies and contextual mismatches even before explicit rule breaches occur. The further development of the empirical literature is based on that premise.

One major theme concerns sequential and streaming detection. Jurgovsky *et al*. proved that recurrent architectures can model order and time dependence of the transactions in a more efficient way compared to the feature sets, which are based on the isolated transactions [Jurgovsky, J. *et al*., 2018]. Other publications like Carcillo *et al*. have demonstrated that unsupervised and supervised learning are applicable particularly in situations where delayed labels, and varying trends of fraud are a drawback in more strictly supervised forms of learning [Carcillo, F. *et al*., 2021]. Also added to the difficulty is the streaming environments where the risk signals need to be interpreted within the constraints of the latencies and partially verified. Under such circumstances the practicality of one model as compared to another is not only based on the accuracy in batch conditions but also based on the ability of the model to provide early warning, stability against drift and the ability to integrate with the operational reviews queues. In this field, the literature continues to suggest that adaptive learning, sliding windows and event history representations are critical because of their ability to maintain behaviour continuity, as opposed to the flattening of conduct into a set of static aggregates.

A second major theme concerns asymmetric error costs and class imbalance. The number of fraudulent or noncompliant events tends to represent a very small proportion of records observed, but instances of missed cases tend to have outsize financial and regulatory cost. Randhawa *et al*. investigated the principles of ensemble learning under these circumstances and reported improved discriminatory ability through the use of boosted and voting-based combinations [Randhawa, K., Loo, C. K. *et al*., 2018]. Makki *et al*. compared the imbalanced learning strategies and discovered that the resampling and calibration of threshold have a significant impact on precision-recall trade-offs [Makki, S. *et al*., 2019]. This research is consistent with more general results of the literature on anomaly detection, which observe that rare-event issues require considerable caution in the selection of metrics since even when a model is inaccurate on the most important cases, the error rate may still be misleadingly high [Ahmed, M. *et al*., 2016]. The risk profiling implication is obvious: the behavioural AI systems cannot be meaningfully assessed without cost-related measures, sensitivity to minority classes, and an explicit comment on alert load.

The third strand is based on relational and network-aware behaviour. Van Vlasselaer *et al*. postulated that the risk is usually not only in the characteristics of the transaction but in the changing relationship between cards, merchants, accounts, and counterparties [Van Vlasselaer, V. *et al*., 2015]. This is particularly applicable in compliance environments that deal with money movement, collusion, mule networks or repetitive low-value transactions with the aim of circumventing thresholds. Network-aware models extend behavioural interpretation to identify interaction topology and, consequently, identify structurally suspicious behaviour that might seem quite normal in a single-transaction setting. These arguments were generalized by Carminati *et al*. who demonstrated that scalable real-time systems could combine a variety of signals to detect malicious behaviour at the production scale [Carminati, M. *et al*., 2015]. Throughout this literature, one can find the idea of graph and rule



Publisher: SARC Publisher

hybridization helpful since the institution-specific knowledge may be used to define the interpretation of suspicious links without giving up the machine-based adaptability.

Behavioural biometrics presents a different yet similar point of view. Teh *et al*. studied the keystroke dynamics and demonstrated that patterns of human-computer interaction can be used as continuous identity information as opposed to authentication credentials that are used once [Teh, P. S. *et al*., 2013]. In the case of compliance remediation systems, this is important since the misuse of accounts and social engineering tend to use valid credentials. Different timing patterns, navigation or input rhythm can therefore be utilized to create a behavioural profile to aid in dynamic authentication, step-up verification or risk-based transaction interruption. Behavioural biometrics offers a risk perspective that is user-centric as compared to transaction only analytics. However, the literature also identifies important limitations such as intra-user variability, sensor dependency, template aging, sensitivity to privacy, and lower cross-device/interfaces transferability. These restrictions warn against the use of behavioural biometrics as an independent control; it seems to have the most significant presence in multimodal systems combining identity, transaction, and contextual evidence.

Another recurring theme is interpretability and governance. Perols demonstrated that the choice of the algorithm has a material impact on the results of fraud detection, but the utility of the model used in the audit environment is also influenced by the explainability and transparency of model outputs [Perols, J. 2011]. The interpretability debate was subsequently extended by Adadi and Berrada to a more methodological level, who stated that even when black-box success is achieved without a justifiable way of reasoning, this can become unacceptable in high-stakes tasks [Adadi, A., & Berrada, M. *et al*., 2018]. Bussmann *et al*. focused specifically on fintech risk management by highlighting that explainable AI can help build managerial trust, regulatory defensibility and enhance model output in relation to risk appetite statements [Bussmann, N. *et al*., 2020]. This literature is especially applicable in behavioural compliance systems since model reasoning often relies on complex temporal or latent characteristics that remain to be transformed into narratives of cases by the analyst. Given developments in reported studies, the literature increasingly suggests that AI scoring relies on the quality of local explanations, provenance of features, and records of decisions.

The literature also reveals several persistent tensions. Deep or ensemble models are likely to perform better than simple baselines but tend to show reduced performance under external validation or rapidly changing attack conditions [Jurgovsky, J. *et al*., 2018; Fiore, U. *et al*., 2019]. Unsupervised/semi-supervised approaches are promising for detecting novelty; however, a high false-positive rate may saturate analysts and ruin remediation capability [Ahmed, M. *et al*., 2016; Carcillo, F. *et al*., 2021]. Richer behavioural data can improve discrimination, whereas the more intense the monitoring, the more privacy, fairness and proportionality issues appear [Mehrabi, N. *et al*., 2021; Rieke, N. *et al*., 2020]. The diversity of methodology has thus been able to bring about partial convergence in progress. There is now solid evidence in the field that analysis of behavioural patterns is more effective in risk identification, although there is less consensus on generalizable architectures, benchmark design, integration of remediation policy, and standards of governance. Table 1 presents the overview of the representative research that informs the existing knowledge on AI-based behavioural risk profiling and associated compliance functions.

**Table 1.** Summary of key findings

| Ref | Focus | Key Findings |
|---|---|---|
| [Jurgovsky, J. *et al*., 2018] | Sequence learning for credit-card transaction streams | Recurrent sequence classification improved detection of temporally patterned fraud relative to transaction-isolated baselines, especially when inter-event order carried signal. |
| [Fiore, U. *et al*., 2019] | GAN-assisted handling of minority fraud cases | Synthetic minority generation improved classifier discrimination under strong imbalance, though robustness depended on distributional fidelity of generated samples. |
| [Carcillo, F. *et al*., 2021] | Combined supervised and unsupervised fraud | Hybrid learning improved adaptability to delayed labels and emerging fraud forms, but increased complexity in |





Publisher: SARC Publisher

| | | |
|---|---|---|
| | detection | explanation and governance. |
| [Teh, P. S. *et al*., 2013] | Keystroke dynamics for continuous behavioural identity assessment | Typing cadence and dwell-flight timing patterns offered persistent identity evidence suitable for continuous authentication and risk-triggered verification. |
| [Perols, J. 2011] | Financial statement fraud analytics | Comparative modelling showed machine learning can outperform traditional statistical procedures, but interpretability and domain specificity remain decisive. |
| [West, J., & Bhattacharya, M. 2016] | Intelligent financial fraud detection review | Broad review identified recurring issues of rarity, drift, feature engineering burden, and lack of standardized evaluation across fraud domains. |
| [Abdallah, A. *et al*., 2016] | Fraud detection systems survey | Survey highlighted imbalance, streaming constraints, concept drift, and the need for adaptive learning architectures in operational fraud systems. |
| [Carminati, M. *et al*., 2015] | Real-time online banking fraud detection architecture | Multi-signal scalable detection supported production-grade alerting, indicating operational feasibility of intelligent monitoring in live banking environments. |
| [Randhawa, K., Loo, C. K. *et al*., 2018] | Ensemble learning for card fraud detection | AdaBoost and majority voting improved classification strength on benchmark fraud data, though threshold tuning remained necessary for deployment realism. |
| [Makki, S. *et al*., 2019] | Imbalanced-learning experiments in fraud classification | Comparative experiments showed that resampling strategy, learner choice, and decision threshold materially alter precision-recall balance. |

Table 1 shows that most studies focus on detection accuracy, with limited attention to remediation design. To a large extent, much of the literature stops at the alert generation level as opposed to following through on how the alerts need to be triaged, explained, and transformed into commensurate corrective action. This is a particularly significant gap when it comes to compliance systems since actionability, defensibility, and burden of review dictate the difference between a high-performing model contributing to institutional control and simply redistributing workload downstream. The following part of the paper thus discusses the conceptual and methodological paradigms which shape the field.

## CONCEPTUAL FRAMEWORK

The literature suggests that AI-based risk profiling systems based on behavioural pattern analysis could be seen as layered socio-technical systems, but not as discrete deployments of models. The majority of reported architectures start with multi-source data capture, then move to feature extraction or representation learning, proceed to risk estimation, and end with some form of analyst review or automated intervention. The strongest studies treat behaviour as a dynamic object. Behaviour is represented through trajectories, temporal patterns, interactions, and recurring events as opposed to the characteristics of a subject which are fixed [Jurgovsky, J. *et al*., 2018; Carcillo, F. *et al*., 2021; Van Vlasselaer, V. *et al*., 2015]. This shift is important. Risk is now perceived as a function of observed behaviour in context, rather than merely of demographic category or account metadata. In compliance remediation, this difference is important since intervention must be based on developing evidence and proportionality, and not merely on historical category identification.

Methodologically, anomaly detection remains one of the field's traditional approaches because it is well suited to rare, novel, and partially labelled forms of noncompliant behaviour. Noncompliant behaviour is often rare, novel, and only partially labelled, hence unsupervised and semi-supervised detection is attractive [Bolton, R. J., & Hand, D. J. 2002; Ahmed, M. *et al*., 2016]. These techniques acquire a representation of how things should be and indicate that anything that is not so is potentially suspicious. The strength of this approach lies in its ability to detect novel cases in which historically labelled fraud cases do not apply to new strategies. This is limited by the fact that anomalous behaviour is not necessarily noncompliant behaviour. A legitimate customer displaying anomalous behaviour can cause a recurrent alert especially when it comes to travelling, expansion of business, replacement of the device or burst of transactions which happens during the season. To contextualize anomaly scores, contextual enrichment and post-detection



Publisher: SARC Publisher

logic is thus needed in compliance practice. In the absence of this enrichment, the systems that are driven by anomalies can cause risk inflation which undermines trust among the analysts.

A second approach to supervised learning is especially used in settings that have a large amount of historical labels and some form of stable feedback. The comparative studies involving the decision trees, logistic regression, support vector machine, ensembles, and deep networks tend to reveal that learned discriminative models ought to outperform simple rules when feature engineering addresses behavioural histories and context [Perols, J. 2011; Randhawa, K., Loo, C. K. *et al*., 2018; Makki, S. *et al*., 2019]. Supervised approaches offer better optimization goals than unsupervised detectors although they have confirmation delay and label bias. Numerous fraudulent or noncompliant incidents cannot be definitively proven or disproven and institutional labels may be grounded on the historical design of rules, rather than ground truth. As a result, purely supervised systems run the risk of becoming historically conservative: good at detecting what past measures of control have already detected, bad at detecting new strategies. This drawback is somehow resolved through the hybrid designs that also involve the application of the exploration based on anomalies and supervised refinement [Carcillo, F. *et al*., 2021].

Sequence models represent a particularly powerful methodological approach since the analysis of behavioural patterns is a time-dependent one. Recurrent neural networks and related sequence models more naturally capture ordered dependence between events, inter-arrival structure, and state transitions than tabular learners do [Jurgovsky, J. *et al*., 2018]. These models are applicable in cases where risk is caused by tendencies like quick merchant switching, frequent near-threshold transfers, alternating device signatures or increasing friction to a payout attempt. Ideally, sequence learning transforms the event-scoring-based compliance monitoring system to trajectory-based interpretation. However, temporal models may be non-transparent, data-intensive and hard to interpret. This limitation is particularly important in regulated applications. Even a high-quality risk score must still be explained clearly in terms of why a sequence is concerning and which remediation step is justified. This has promoted an increase in the interest in local explanation methods and surrogate summary mechanisms, but in practice they are not yet fully mature [Adadi, A., & Berrada, M. *et al*., 2018; Bussmann, N. *et al*., 2020].

Graph and network methods extend temporal logic into relational space. It can be said that financial crime and policy circumvention are usually based on related participants, reuse of beneficiaries, partnering of merchants, and indirect transaction routes. Graph-based models represent structural properties such as proximity, centrality, community structure, and anomalous link formation [Van Vlasselaer, V. *et al*., 2015]. In theory, this mechanism would be in line with the compliance issues with regard to clandestine coordination and decentralized malpractices. A major strength of graph-based methods is their ability to reveal the collective behaviour that is not visible in single record-at-a-time screenings. Its main weakness lies in operational complexity: the choices of graph construction, the latency of updates, sparse labels, and difficulty of explanations can all reduce deployability. Graph methods have been found to be more effective in most institutions as an enrichment layer that facilitates case escalation as opposed to being a single decision engine.

Another difference in methodology is with respect to cost-sensitive and imbalance-conscious learning. Due to the huge disparities between the costs of false negatives and false positives, thresholding and class weighting are central design choices rather than minor tuning parameters [Randhawa, K., Loo, C. K. *et al*., 2018], [Makki, S. *et al*., 2019]. It has been reported that when evaluation shifts from global accuracy to precision-recall measures, area under the precision-recall curve, cost curves or alert-rate constrained recall, the performance metrics change significantly. Such transition has direct compliance importance. A smaller investigative institution may prefer a narrower alert stream, or an institution with high regulatory pressure may tolerate a higher false-positive rate to cut down on missed suspicious cases. The comparison of methods in the literature, thus, shows that there is something deeper: the technical superiority depends on the capacity of remediation, tolerance of the policy, and operational economics.

Methodological choice is also determined by interpretability and fairness. The explainable AI research has postulated that the choice of models to use in high-stakes areas should be informed by comprehensibility, contestability and governable fit and not just predictive strength [Adadi, A., &



Publisher: SARC Publisher

Berrada, M. *et al*., 2018; Bussmann, N. *et al*., 2020]. Fairness literature provides a warning against the unequal distribution of errors among the groups of customers, regions, behaviour patterns or device access time patterns [Mehrabi, N. *et al*., 2021]. Behavioural systems are particularly susceptible since the data of conduct can reflect a socio-economic situation, access limitations, language, or digital literacy. A strict deviation model may therefore penalise legitimate users who may not be standard. Methodological strategies that incorporate explanation, drift control, confidence assessment, and calibrated intervention levels are therefore better suited to compliance remediation as compared to architectures optimized solely for benchmark gains.

Figure 1 presents a conceptual framework that organizes the literature into a risk-to-remediation behavioural pipeline. The diagram demonstrates that raw behavioural traces are converted into features or learned representations, multiple analytic engines generate risk evidence, and outputs are converted to compliance action. The diagram promotes the main idea of this review that detection quality and remediation quality are two parts of the same mechanism and not two stages.

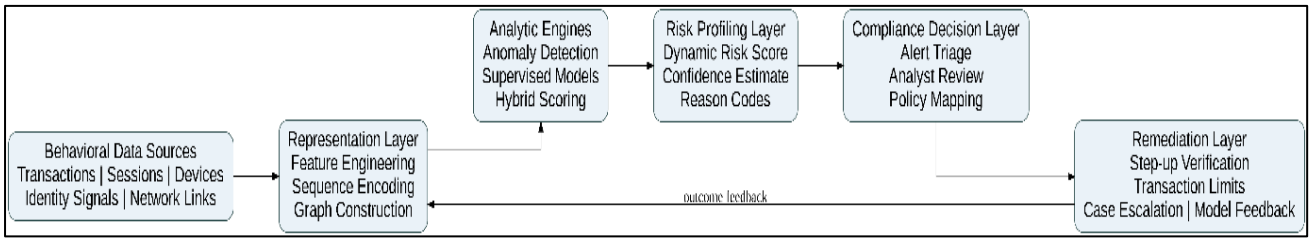


**Figure 1.** Conceptual framework for AI-driven behavioural risk profiling and compliance remediation

It can be concluded that the methodological literature is oriented to a composite design principle. Effective systems incorporate temporal, relational, and contextual behavioural features. They are adaptive, uncertainty-aware, explainable, and responsive to remediation feedback. Weaker approaches isolate one component, such as high-capacity classification, without integrating it into the broader operational framework required by compliance work.

## DISCUSSION

According to reported results in the literature, behavioural pattern analysis tends to be much more effective than static or rule-based risk discrimination schemes, but the magnitude and consistency of the improvement depend extensively on the data regime and deployment assumptions. A significant improvement in sequence-aware models is frequently measurable due to suspicious strategies being captured by transaction order, timing and local context which are washed out by feature aggregation [Jurgovsky, J. *et al*., 2018]. Stronger minority-class detection has also been reported with GAN-assisted and ensemble approaches with severe imbalance [Fiore, U. *et al*., 2019; Randhawa, K., Loo, C. K. *et al*., 2018]. However, such benefits are not uniform. The performance uplift in certain environments is extremely sensitive to preprocessing decisions, label quality and the realism of the evaluation design and data split. Random train-test partition studies can be found to overestimate performance in the event that there is temporal leakage or repeated actor patterns across partitions. These findings support the usefulness of behavioural AI, albeit with a warning against the blind comparison of headline measures across unrelated datasets.

One consistent finding is the importance of hybridization. According to Carcillo *et al*., the combination of both supervised and unsupervised aspects leads to benefits particularly when confirmed labels are delayed and new attack patterns occur during the lifecycle of the models [Carcillo, F. *et al*., 2021]. The implication of this finding is highly relevant to compliance operations since an institution can seldom work with perfect, complete, and unbiased labels. A hybrid system is able to utilize the past knowledge and still be sensitive to new deviations. Nevertheless, as it is demonstrated in the literature, hybridization introduces governance complexity. A supervised score that is based on historical cases may be easier to comprehend by an analyst as compared to an anomaly score that is based on latent deviation. In the case of fusing such signals, the quality of explanation is of critical importance. A remediation response based on the presence of opaque score fusion can be more difficult to defend than a less accurate workflow which is more interpretable. Although hybrid systems are technically attractive, their institutional value



**Publisher: SARC Publisher**

depends on explanation quality and case-management integration.

The pattern of results of behavioural biometrics is different. The interaction dynamics and keystroke patterns may provide a form of continuous assurance in addition to passwords and one-time authentication that supports risk-based verification during suspicious sessions [Teh, P. S. *et al*., 2013]. From a compliance perspective, this matters since identity compromise is one of the common precursors of unauthorized transfers, policy violations or abuse of allowed privileges. Behavioural biometrics is hence a preventative layer, which minimizes downstream remediation load since account abuse is detected earlier in the session. Nonetheless, the reported outcomes are sensitive to the environment drift, interface drift, fatigue, accessibility requirements as well as device drift. This kind of sensitivity restricts the direct portability to universal compliance infrastructures. Better results are observed in cases where the behavioural biometrics is combined with transaction and contextual indicators instead of using it as a conclusive identity indicator. This observation is consistent with a more general finding of the literature: no single behavioural modality is sufficient across heterogeneous institutional settings.

Another major empirical constraint identified in the reviewed studies is class imbalance. Makki *et al*. demonstrated that the model ranking can change significantly based on the resampling strategy and the metric used [Makki, S. *et al*., 2019]. Randhawa *et al*. discovered that ensemble techniques are useful for detection, but deployment value depends on the calibration of threshold and control of alert rates [Randhawa, K., Loo, C. K. *et al*., 2018]. This is important since many institutional evaluations continue to rely on the area under the ROC curve as a primary performance measure. Such a practice is poor in the rare-event compliance settings. Precision, recall, the number of false-positives, alert queue size, and investigator throughput is more operationally informative. The literature is very strong in indicating that behavioural risk profiling must be considered against the institutional constraints and not against the abstract classification objectives. The detector which increases the recall by a small percentage can be helpful in one setting and operationally harmful in the other depending on the remediation ability and impact tolerance of the customers.

In the reviewed literature, interpretability is determined as an enabling as well as a limiting factor. Explainable AI studies indicate that the area of application in risk-sensitive sectors enhances the adoption process when analysts are able to review the reason codes, feature attribution, and confidence indicators [Adadi, A., & Berrada, M. *et al*., 2018; Bussmann, N. *et al*., 2020]. Practically, interpretability has a greater impact than trust; it has an effect on the remediation accuracy. The inclusion of an alert that includes an abnormal beneficiary sequence, a change of device, and a spike in nighttime velocity, implies that the intervention required is quite different than an alert that displays a generic high-risk score. The former can make step-up verification or temporary transfer friction to be implemented selectively, the latter can result in the effect of blunt account restriction. According to reported studies, a more granular explanation can be supported by a more proportionate remediation, which, in turn, can reduce customer harm and case backlog. Simultaneously, a simplified description can distort the way in which a complex model came up with a decision. This creates a tension between faithful explanation and operational usability, which has not been resolved in most of the literature.

Another issue that makes results interpretation hard is fairness and governance issues. Behavioural models have unintended effects whereby atypical but legitimate behaviour is mapped to higher risk in cases where the training data are skewed by uneven reporting, increased surveillance or past intervention bias [Mehrabi, N. *et al*., 2021]. As an illustration, the behaviour of being out of place at odd times of the day, unstable access to the devices, or frequent change in location can be associated with valid occupational or socio-economic conditions but not with malpractice. A compliance remediation system that intensifies scrutiny in all such cases may disproportionately burden certain user groups. Based on the reported studies, subgroup analysis is rarely reported, constituting a major gap. Excellent results in general performance will not ensure fair intervention patterns. It is among the reasons why fairness auditing, calibration tests, and safeguards of the human review should receive greater attention in future empirical work. With compliance systems, legitimacy is not only based on the quality of detection but also how the friction and suspicion is distributed.



**Publisher: SARC Publisher**

The other area where the literature is oriented towards the future usefulness is privacy-preserving deployment. Although federated learning studies have been produced in related fields, they can provide methodological knowledge on how sensitive behavioural information could be used to help train a common model without much centralization [Rieke, N. *et al.*, 2020]. In compliance scenarios where cross-branch, cross-subsidiary or consortium-based risk detection is to be used, such techniques may minimize data transfer and yet retain some collaborative learning advantage. However, privacy-preserving design can also conflict with explainability and auditability. Distributed training may also make it more difficult to trace model behaviour, and the privacy requirements may also restrict the granularity of evidence that can be used to construct remediation narratives. The literature hence indicates that the concept of privacy enhancement is not a simple add-on. It transforms the informational form of risk profiling, and it might necessitate new governance systems to explain and deal with challenges.

The following three tables compare approaches and reported results of representative studies. Table 2 dwells on the strengths and weaknesses of the methodology. Table 3 puts emphasis on systems, measures and outcomes that have been reported in the literature. These comparisons combined demonstrate that AI-based behavioural methods are effective under some conditions and less effective under others, and it is impossible to have a family of methods that can be effective in all compliance-related factors.

**Table 2.** Method comparison

| Ref | Method | Strengths | Limitations |
|---|---|---|---|
| [Jurgovsky, J. *et al*., 2018] | Recurrent sequence classification | Captures temporal order, inter-transaction dependence, and evolving local context. | Opaque internal states, higher data demand, explanation difficulty in regulated settings. |
| [Fiore, U. *et al*., 2019] | GAN-assisted minority augmentation | Improves learning under severe imbalance and sparse fraud labels. | Sensitive to synthetic sample realism; can amplify artifacts if minority distribution is poorly represented. |
| [Carcillo, F. *et al*., 2021] | Hybrid supervised-unsupervised scoring | Balances known-pattern learning with novelty sensitivity and delayed-label resilience. | Score fusion and alert explanation become more complex for analysts and governance teams. |
| [Teh, P. S. *et al*., 2013] | Keystroke dynamics biometrics | Supports continuous identity assurance and session-level anomaly detection. | Device dependency, template aging, accessibility variation, privacy concerns. |
| [Perols, J. 2011] | Comparative statistical and machine learning fraud models | Supports model transparency comparison and domain-grounded feature interpretation. | May underperform on nonlinear temporal patterns without richer behavioural representation. |
| [Carminati, M. *et al*., 2015] | Scalable real-time banking fraud engine | Operationally feasible, multi-signal, and suitable for low-latency decision environments. | Institution-specific engineering burden and potential difficulty in transfer across platforms. |
| [Randhawa, K., Loo, C. K. *et al*., 2018] | Ensemble boosting and majority voting | Strong discrimination on benchmark card-fraud data and robustness across learners. | Threshold management remains critical; benchmark performance may not map directly to live settings. |
| [Makki, S. *et al*., 2019] | Imbalance-aware experimental comparison | Clarifies metric sensitivity and learner dependence under rare-event conditions. | Results remain dataset contingent; comparative rankings shift with resampling choice. |
| [Van Vlasselaer, V. *et al*., 2015] | Network-based fraud detection | Captures collusion, shared infrastructure, and suspicious relational topology. | Graph maintenance cost, sparse labels, and complex explanation pathways. |





**Publisher: SARC Publisher**

**Table 3.** Results comparison

| Ref | System | Metric | Outcome |
|---|---|---|---|
| [Jurgovsky, J. *et al*., 2018] | Sequence-based card transaction detector | Precision-recall performance | Temporal modelling improved detection of fraud sequences over transaction-isolated baselines. |
| [Fiore, U. *et al*., 2019] | GAN-enhanced fraud classifier | Classification effectiveness | Minority augmentation improved classifier performance on imbalanced card-fraud data. |
| [Carcillo, F. *et al*., 2021] | Hybrid fraud detection framework | Detection adaptability | Combined learning improved resilience to delayed labels and emerging fraud behaviour. |
| [Teh, P. S. *et al*., 2013] | Continuous behavioural authentication system | Authentication accuracy | Typing-pattern models provided persistent user differentiation suitable for ongoing risk checks. |
| [Perols, J. 2011] | Financial statement fraud detector | Comparative predictive accuracy | Machine learning models outperformed several traditional approaches on fraud identification tasks. |
| [Carminati, M. *et al*., 2015] | Online banking fraud monitoring platform | Real-time detection capability | Production-oriented system supported scalable intelligent detection in operational banking. |
| [Randhawa, K., Loo, C. K. *et al*., 2018] | AdaBoost and voting-based fraud detector | Classification performance | Ensemble structure improved discriminatory power relative to several single-model baselines. |
| [Makki, S. *et al*., 2019] | Imbalanced fraud classification pipeline | Precision-recall balance | Performance varied substantially with sampling and threshold strategy, underscoring deployment sensitivity. |
| [Van Vlasselaer, V. *et al*., 2015] | APATE network-based fraud detector | Fraud detection effectiveness | Relational features improved detection of suspicious card activity beyond isolated transaction features. |

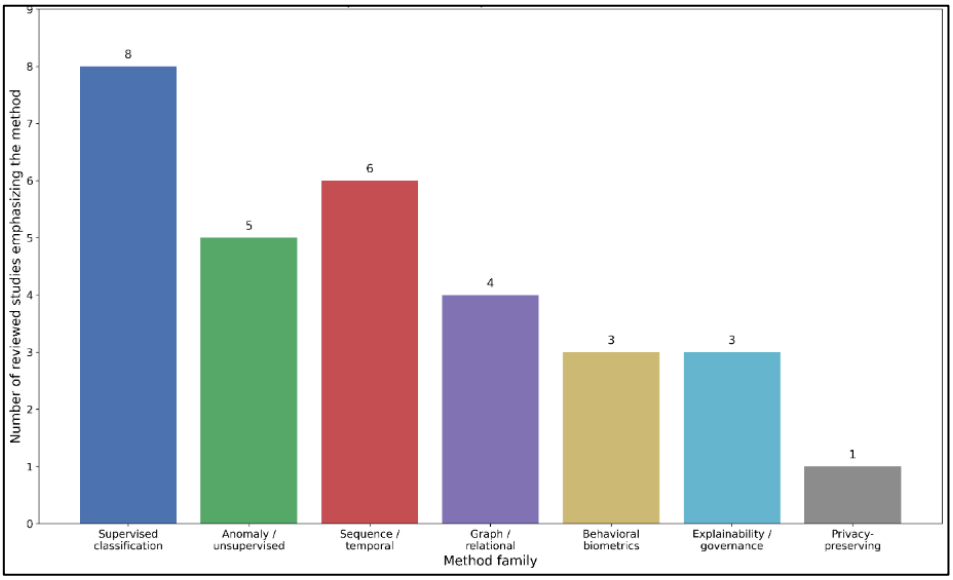


**Figure 2.** Thematic distribution of methodological emphasis across representative reviewed studies

Figure 2 converts the literature into a bare thematic trend graph with the number of counts of emphasis on methods across the representative studies in Tables 1 to 3. The graph does not purport to have a bibliometric census. Rather, it visualizes the intensity of the various methodological families that are evident in the reviewed evidence base. The trend indicates that there is high concentration in the use of supervised, imbalance-conscious and time-constrained approaches, and relatively less emphasis, though growing emphasis, on explainability and privacy-preserving design.





**Publisher: SARC Publisher**

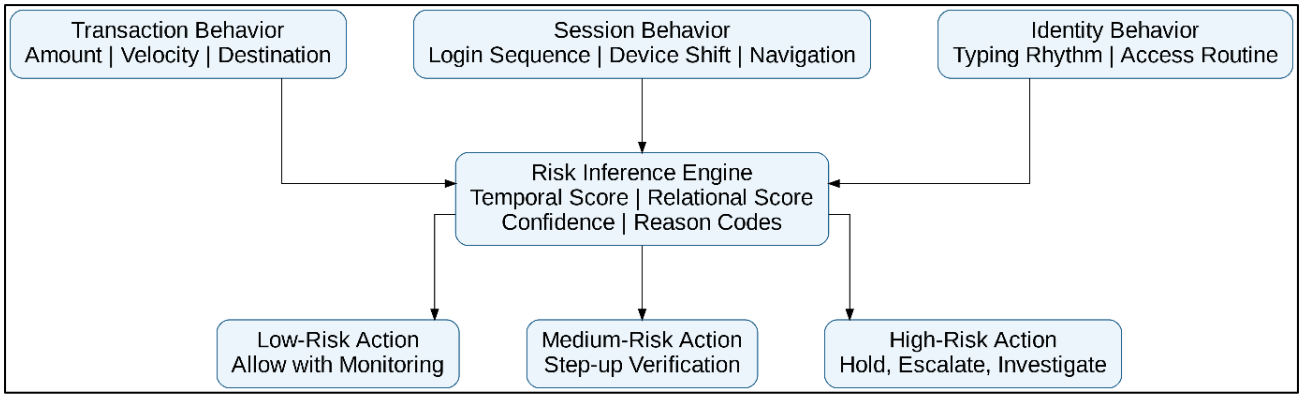


**Figure 3.** Relationship diagram between behavioural signals, inferred risk, and compliance action

The reasoning that is summarized in Figure 3 explains why compliance remediation should be given more consideration in the literature. A behavioural cue has limited institutional value until it is linked to an appropriate response pathway. An overly blunt mapping is more frictional and expensive, whereas a more liberal mapping is less protective.

Figure 4 further illustrates an integrated deployment and governance model. The diagram incorporates the feedback of investigator outcome and remediation result on the data curation, threshold tuning, and model monitoring. This closed loop view is indicative of a recurring finding in the literature: the current deployment paradigm is ill-suited to changing behavioural contexts.

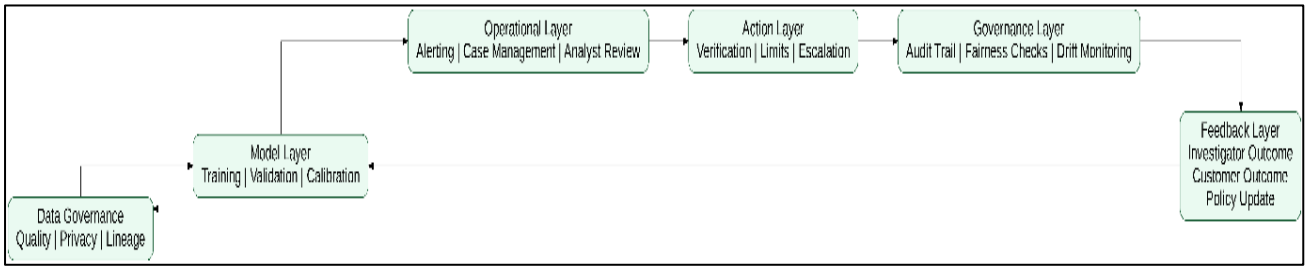


**Figure 4.** Integrated model for deployment, remediation, and governance

The structure in Figure 4 is integrated to strengthen the overall conclusion of the article. Behavioural AI is only capable of fostering greater compliance results when model outputs, remediation decisions, oversight checks, and feedback learning are implemented as a systemic control architecture. Lacking governance in detection creates a weak benefit; lacking adaptive behavioural intelligence in governance causes institutions to be vulnerable to emerging malpractice.

## FUTURE DIRECTIONS

One of the main directions future research should take is evaluation realism. Future research should evaluate models using temporally ordered validation instead of random splits. Studies should also report alert-capacity constraints, confirmation delays, intervention costs, and subgroup performance. External validity also requires more work to be done. Models that work well in one stream of payment, one geography or one segment of customers can fall drastically in another due to the differences in behavioural norms. The empirical agenda would be stricter in the event that it incorporates multi-site evaluation, drift-resistant benchmarking as well as real reporting of truthful load per true positive. Such an activity has the potential of taking the field out of the alleged small performance to the institutionally pertinent evidence.

A second priority is the development of explanation architectures that are adapted to the compliance operations. The current explainability techniques normally present the scores of feature importance or local approximations, though, to correct, an expedient approach is required other than the post hoc explanation. The researchers need clear behavioural descriptions, level of uncertainty and courses of action attributed to alternatives in the policy. Future studies should examine the forms of explanation supporting the differentiation of action, e.g., enhanced due diligence, temporary friction, soft outreach, and case closure. It would particularly be important to



Publisher: SARC Publisher

compare the inherently interpretable temporal models with black-box architectures augmented by explanation overlays. The advancements in this field may minimize the difference in time between the technical detection and defensible compliance action.

A third orientation involves fairness, proportionality as well as privacy. The strength of behavioural data lies in its ability to capture fine-grained behaviour, but this granularity creates risks of over-surveillance and unequal treatment. The work of the future must consider fairness auditing as a part of model development and post-deployment monitoring and concentrate on false-positive concentration and the remediation intensity of user groups [Mehrabi, N. *et al*., 2021]. Greater attention should be paid to privacy-preserving collaborative learning, in particular, to federated or distributed risk profiling, in which sensitive data cannot be pooled at will [Rieke, N. *et al*., 2020]. Nevertheless, the enhancement of privacy should not be treated as an add-on in future studies. It should also be given the same consideration as to its effect on explainability, audit trails, and challenge rights.

A fourth research requirement is closed-loop remediation science. The current literature pays substantial attention to the scoring of suspicious behaviour, but there is a relative lack of attention paid to the investigation of the most effective remedial responses, the least disruptive, and fairest in uncertain situations. Future research needs to look at intervention design in an experimental way: step-up verification timing, graduated friction, dynamic limits, routing logic of cases, and incorporation of feedback based on the results of the analysts. This would encourage the field to treat compliance systems as adaptive control mechanisms rather than discrete detectors. This area is particularly promising for interdisciplinary collaboration with compliance specialists, human factors scholars, legal analysts, and operations experts in the next round of development.

## CONCLUSION

The reviewed literature shows that AI-based risk profiling grounded in behavioural pattern analysis has become an important and increasingly sophisticated field of study. In the context of fraud detection, online banking security, anomaly detection, and behavioural authentication, reported studies generally indicate that temporal, contextual, and relational cues enhance the capacity to identify suspicious or noncompliant behaviour as compared to the application of fixed rules as a profiling measure. Hybrid, sequence models, graph-based, and behaviour-aware classifiers are all significant especially when it comes to data environments such as rare, drift and adversarial adaptation.

In the meantime, there are certain definite limits of the literature. Handling of imbalance, design of validation, quality of label and deployment scenario is still very sensitive to performance. Four areas, interpretability, fairness, and privacy, and governance are not peripheral concerns, but the main conditions under which a behavioural AI system can be said to be able to support legitimate compliance action. The alert generating process is not adequate. The bigger challenge lies in translating indications of hazards into proportionate, clarifiable and functioning way of remedies. This is one of the gaps that need to be narrowed to implement AI systems under controlled environments.

The new conception of detection and remediation, thus, is an academic value. Any additional development will not be grounded on any particular metric improvement but on architectures of systems that will take into account behavioural intelligence, quality of explanations, operational fit and discipline governing the architecture. The area remains a prospect, but, a sustainable impact will depend on careful institutional planning, a tight governance, and a stronger emphasis on research about feedback-based remediation.

## REFERENCES

1. Ngai, E. W., Hu, Y., Wong, Y. H., Chen, Y., & Sun, X. "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature." *Decision support systems* 50.3 (2011): 559-569.
2. Bolton, R. J., & Hand, D. J. "Statistical fraud detection: A review." *Statistical science* 17.3 (2002): 235-255.
3. Ahmed, M., Mahmood, A. N., & Islam, M. R. "A survey of anomaly detection techniques in financial domain." *Future Generation Computer Systems* 55 (2016): 278-288.
4. Adadi, A., & Berrada, M. "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)." *IEEE access* 6 (2018): 52138-52160.
5. Bussmann, N., Giudici, P., Marinelli, D., & Papenbrock, J. "Explainable AI in fintech risk





**Publisher: SARC Publisher**

management." *Frontiers in Artificial Intelligence* 3 (2020): 26.
6. Jurgovsky, J., Granitzer, M., Ziegler, K., Calabretto, S., Portier, P. E., He-Guelton, L., & Caelen, O. "Sequence classification for credit-card fraud detection." *Expert systems with applications* 100 (2018): 234-245.
7. Fiore, U., De Santis, A., Perla, F., Zanetti, P., & Palmieri, F. "Using generative adversarial networks for improving classification effectiveness in credit card fraud detection." *Information Sciences* 479 (2019): 448-455.
8. Carcillo, F., Le Borgne, Y. A., Caelen, O., Kessaci, Y., Oblé, F., & Bontempi, G. "Combining unsupervised and supervised learning in credit card fraud detection." *Information sciences* 557 (2021): 317-331.
9. Teh, P. S., Teoh, A. B. J., & Yue, S. "A survey of keystroke dynamics biometrics." *The Scientific World Journal* 2013.1 (2013): 408280.
10. Perols, J. "Financial statement fraud detection: An analysis of statistical and machine learning algorithms." *Auditing: A Journal of Practice & Theory* 30.2 (2011): 19-50.
11. West, J., & Bhattacharya, M. "Intelligent financial fraud detection: A comprehensive review." *Computers & security* 57 (2016): 47-66.
12. Abdallah, A., Maarof, M. A., & Zainal, A. "Fraud detection system: A survey." *Journal of Network and Computer Applications* 68 (2016): 90-113.
13. Carminati, M., Caron, R., Maggi, F., Epifani, I., & Zanero, S. "BankSealer: Scalable real-time intelligent fraud detection in online banking." Computers & Security, 48.1 (2015): 175–192.
14. Randhawa, K., Loo, C. K., Seera, M., Lim, C. P., & Nandi, A. K. "Credit card fraud detection using AdaBoost and majority voting." *IEEE access* 6 (2018): 14277-14284.
15. Makki, S., Assaghir, Z., Taher, Y., Haque, R., Hacid, M. S., & Zeineddine, H. "An experimental study with imbalanced classification approaches for credit card fraud detection." *Ieee Access* 7 (2019): 93010-93022.
16. Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. "A survey on bias and fairness in machine learning." *ACM computing surveys (CSUR)* 54.6 (2021): 1-35.
17. Rieke, N., Hancox, J., Li, W., Milletari, F., Roth, H. R., Albarqouni, S., & Cardoso, M. J. "The future of digital health with federated learning." *NPJ digital medicine* 3.1 (2020): 119.
18. Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., & Baesens, B. "APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions." *Decision support systems* 75 (2015): 38-48.

**Source of support:** Nil; **Conflict of interest:** Nil.

**Cite this article as:**
Vichare, S. C. "AI-Driven Risk Profiling and Compliance Remediation System Based on Behavioural Pattern Analysis." *Sarcouncil Journal of Applied Sciences* 6.4 (2026): pp 19-31.





**Publisher: SARC Publisher**