International Journal of Commerce and Management Research
www.managejournal.com
ISSN: 2455-1627
Received: 10-06-2026, Accepted: 10-07-2026, Published: 25-07-2026
Volume 12, Issue 3, 2026, Page No. 120-126

Risk-based customer monitoring using statistical classification in branch banking

Mohammad Nurul Islam[1]*, Md Rezwan Ur Rahman[2], Rasheda Yasmin[3], Jahin Tajowar Masud[4], Md. Shakil[5]
[1] Department of Statistics, University of Dhaka, Dhaka, Bangladesh
[2] Eastern Bank PLC, Dhaka, Bangladesh
[3] Department of Mathematics, University of Chittagong, Chittagong, Bangladesh
[4] Department of Computer Science and Engineering, Jagannath University, Dhaka, Bangladesh
[5] Department of Law, Premier University, Chittagong, Bangladesh
DOI: https://doi.org/10.66856/ijcmr.2026.12.3.12180

Abstract
Branch banking continues to depend on relationship knowledge, local customer familiarity and routine operational checks, yet many branches struggle to separate ordinary customer variation from early risk signals that require supervisory attention. This paper develops a practical risk-based customer monitoring framework using statistical classification for branch banking environments. The objective is to show how routine customer, account-conduct and transaction indicators can be converted into low, medium and high monitoring bands that support proportionate branch action. The paper is designed as a methodological and conceptual research article rather than a completed empirical study. It synthesizes credit scoring, behavioral monitoring, anomaly detection, model validation and banking governance literature to propose a branch-level classification architecture. The framework uses structured customer variables, account behavior, product exposure, transaction irregularity, service interaction and branch judgment as inputs to a transparent classifier such as logistic regression, discriminant analysis, decision-tree rules or scorecard-based segmentation. The output is not only a predicted risk label but a management response: routine monitoring, enhanced review, documentary verification, supervisory approval or escalation to risk and compliance teams. The paper contributes by translating statistical classification into a bank-operational monitoring process that can be implemented even by branches with moderate data maturity. It also emphasizes calibration, class imbalance, explainability, override documentation and periodic model review so that customer monitoring remains auditable, fair and useful for managers.

Keywords: Branch banking, customer monitoring, statistical classification, risk-based approach, credit scoring, account conduct, risk management, logistic regression, decision rules, bank operations

Introduction

Risk-based customer monitoring is a central requirement of modern branch banking. Banks are expected to know their customers, monitor account behavior, detect changes in risk status and apply proportionate action before losses, compliance breaches or service failures become severe. In practice, however, customer monitoring at the branch level is often uneven. Relationship managers may understand local customers well, but their knowledge is not always recorded in a consistent format. Operational staff may observe unusual account behavior, but those signals can remain fragmented across teller records, transaction systems, overdraft reports, service requests and manual notes. As a result, branch managers may spend considerable time reviewing low-risk cases while some high-risk customers receive attention only after arrears, complaints, irregular transactions or documentation gaps have already escalated.

Statistical classification offers a practical way to improve this process. A classification model assigns observations to defined categories on the basis of measurable characteristics. In branch banking, the categories may be low, medium and high monitoring risk, or routine, enhanced and escalated review. The goal is not to replace branch judgment. Instead, classification can organize scattered information into a disciplined monitoring process that helps managers identify which customers need attention, why they were flagged and what action should follow. When used responsibly, such a system can reduce arbitrary monitoring, improve consistency across branches and create an audit trail for risk and compliance functions.

The need for structured monitoring is especially important in commercial and retail branch networks where data quality and customer behavior vary widely. Branches handle small businesses, salaried customers, micro-enterprises, deposit customers, borrowers, inactive accounts, cash-intensive businesses and digitally active clients. Each group produces different risk signals. For example, a declining deposit balance may be normal for a seasonal small business but concerning for a salary account that recently received an overdraft facility. A rise in transaction volume may indicate business growth, but it may also require verification if it is inconsistent with the customer profile. A risk-based framework should therefore evaluate customer behavior relative to product, sector, account history and documented customer profile.

This paper develops a practical framework for risk-based customer monitoring using statistical classification in branch banking. It is written for a management and commerce audience and therefore focuses on operational use, not only statistical technique. The paper explains the data domains that branches can monitor, how a statistical classifier can be designed, how risk bands can be linked to branch actions and how the system can be validated and governed. Because no bank-level confidential data are available for this manuscript, the paper does not report empirical model accuracy or fabricate performance results. Instead, it proposes a validation protocol that banks may apply to their own branch data.

The contribution of the paper is threefold. First, it reframes customer monitoring as a classification-and-action problem rather than a manual checklist exercise. Second, it integrates customer profile, account conduct, transaction behavior, product exposure and staff judgment into a single monitoring architecture. Third, it emphasizes governance requirements such as calibration, class imbalance, explainability, override tracking and fairness review, which are often ignored when branches adopt simple red-flag tools. The remainder of the paper presents the literature background, framework design, statistical classification method, implementation process, governance controls, limitations and conclusion.

Literature Review

Statistical classification has a long history in banking and finance. Early credit-risk research used financial ratios and discriminant analysis to classify firms according to default risk (Altman, 1968) [3]. Later consumer credit scoring studies adopted logistic regression, scorecards and decision rules because these methods were transparent and easy to operationalize (Hand and Henley, 1997; Thomas, 2000) [14, 17]. More recent studies compare machine-learning algorithms such as random forests, gradient boosting and support vector machines, often reporting stronger predictive performance under certain data conditions (Brown and Mues, 2012; Lessmann *et al*., 2015; Dumitrescu *et al*., 2022) [7, 10, 15]. However, the banking value of a classifier depends not only on its technical accuracy but also on whether the output can be explained, challenged and linked to a decision.

Credit scoring literature is relevant because customer monitoring also requires ranking and classification. Traditional credit scoring estimates the likelihood that a borrower will default and has developed a wide set of evaluation criteria for classification models (Abdou and Pointon, 2011) [1]. Customer monitoring is broader. It may include deterioration in account conduct, unusually rapid transaction changes, excessive cash activity, inactive or dormant behavior, documentation gaps, frequent complaints, missed payments, repeated limit excesses or other indicators that the customer profile should be reviewed. This makes branch monitoring a multi-risk process rather than a single default prediction task.

Behavioral scoring and account-conduct models show that dynamic customer behavior can be more useful than static application data after a customer relationship has begun (Thomas, 2000) [17]. Account turnover, utilization, repayment timing, returned cheques, arrears, deposit volatility and product usage can reveal changes in risk status before formal default or closure. In branch banking, these variables are often available at monthly or even daily frequency. Their usefulness depends on data cleaning, consistent customer identifiers and careful treatment of missing or seasonal records.

Anomaly and fraud detection literature also informs branch monitoring (Bolton and Hand, 2002) [6]. Banks must identify patterns that deviate from expected behavior, but not every deviation is suspicious or risky. A purely rule-based system may generate many false alerts, creating review fatigue among branch staff. A statistical classifier can combine several weak indicators into a more reliable risk category, reducing dependence on single red flags. At the same time, classification models can be affected by class imbalance because serious risk events are usually rare compared with normal customer activity (Brown and Mues, 2012; Chen *et al*., 2024) [7, 9]. This means that simple accuracy is an unsuitable performance measure; precision, recall, capture rate, calibration and workload measures are needed.

The literature on explainable machine learning is particularly relevant to branch banking (Bussmann *et al*., 2021; Chen *et al*., 2024) [8, 9]. A model that cannot explain why a customer was assigned to a high-risk band will be difficult for branch staff to trust and difficult for managers to defend. Local reason codes, variable contribution summaries and clear decision rules can help translate model outputs into practical action. For example, a branch manager should be able to see that a customer was escalated because account turnover fell sharply, overdraft utilization remained high for three months and customer documentation was overdue.

Governance literature further shows that risk models require continuous monitoring and model-risk review (Alonso Robisco and Carbó Martínez, 2022) [2]. A branch classification model may work during stable conditions but perform poorly during inflation, economic slowdown, technological migration, branch restructuring or changes in customer behavior. Therefore, periodic validation, drift analysis, threshold review and override monitoring are necessary. Human judgment remains important, but it should be documented. If branch managers override classifications without evidence, the model may become symbolic rather than effective.

Conceptual Framework for Branch Customer Monitoring

The proposed framework treats customer monitoring as a continuous cycle of data capture, variable coding, statistical classification, risk band assignment, branch action and feedback. The unit of observation can be a customer-month, account-month or customer-review period, depending on the bank system. The preferred unit is customer-month because it supports repeated monitoring and allows managers to identify movement across risk bands over time.

The framework begins with branch customer records. These include customer type, occupation or business activity, account age, product mix, KYC status, credit exposure, deposit profile, transaction activity, service history and prior risk events. The second stage converts these records into monitoring variables. Variables must be measurable, auditable and meaningful. For example, "customer is risky" is not a useful variable, but "three overdraft excesses within 90 days" or "cash deposits exceed declared profile by more than a defined percentage" can be used as a structured feature.

The third stage applies a statistical classifier. The classifier may be a logistic regression model, discriminant classification, decision-tree model, points-based scorecard or hybrid rule-statistical model. The exact technique should depend on available data, event frequency, and interpretability needs and staff capacity. A low-maturity branch network may begin with a scorecard based on documented weights and later compare it with logistic regression. A larger bank may use a challenger model such as gradient boosting, provided explainability and validation controls are in place.

The fourth stage assigns a customer to a risk band. The band must be linked to an action. A low-risk customer continues under routine monitoring. A medium-risk customer may require profile verification, relationship-manager contact or additional document review. A high-risk customer may require branch-manager review, enhanced monitoring, risk department referral or temporary restriction until the concern is resolved. Without a defined action, a classification model becomes only a reporting tool and does not improve branch management.

The fifth stage records feedback. If a customer is classified as high risk but the branch manager finds a legitimate explanation, the override should be recorded with evidence. If a customer later defaults, closes the account, generates a compliance alert or receives a confirmed complaint, that outcome should be used to recalibrate the classifier. Feedback is necessary because customer monitoring is dynamic and branch risk patterns change over time.

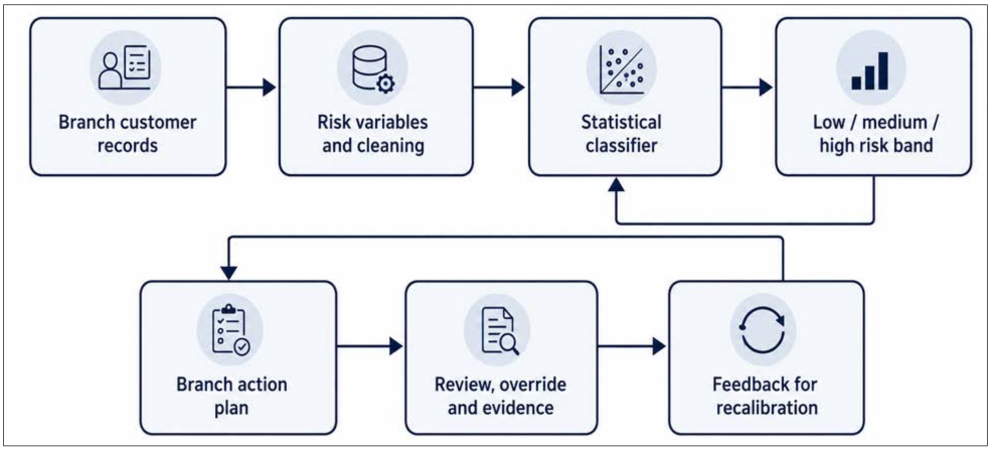

Fig 1: Risk-based customer monitoring cycle in branch banking

Data Domains and Monitoring Indicators

A practical classification system should use variables that are already available or can be collected without excessive burden. The proposed data domains are customer profile, account conduct, transaction behavior, product exposure, service interaction and branch judgment. These domains are broad enough for retail and small-business branch portfolios, yet specific enough to support statistical classification.

Customer profile variables describe the baseline expectation for behavior. These include customer segment, occupation or business type, account age, geographic location, declared income or turnover, expected transaction range and documentation status. Account conduct variables capture how the customer actually uses the account. These include balance volatility, deposit turnover, frequency of low-balance days, overdraft usage, cheque returns, missed instalments and dormant behavior. Transaction behavior variables capture changes in volume, value, frequency, cash intensity, inward and outward transfers and unusual spikes relative to the customer's own history.

Product exposure variables are important because monitoring risk increases with product complexity. A customer with only a savings account may require different monitoring than a customer with a current account, trade facility, overdraft, credit card and business loan. Service interaction variables include complaints, repeated requests for limit changes, delayed document submission, account closure threats and unfulfilled branch commitments. Finally, branch judgment converts staff observations into coded inputs. For example, a site-visit concern, management dispute, sudden change in business activity or repeated avoidance of contact can be recorded using defined categories.

The key requirement is standardization. Branch staff should not enter free-text risk opinions without coding rules. Free text can be valuable, but it should support evidence rather than become the only monitoring input. A bank may use standardized reason codes such as documentation overdue, unexplained turnover increase, repeated limit excess, account inactivity after loan disbursement, or customer unreachable. These codes allow analysis across branches and reduce inconsistency.

Table 1: Data domains for risk-based customer monitoring in branch banking

Data domain	Example indicators	Monitoring purpose
Customer profile	Customer type, occupation/business activity, account age, location, declared income or turnover, expected transaction range, KYC/document status.	Defines expected behavior and baseline risk context.
Account conduct	Balance volatility, deposit turnover, overdraft usage, cheque returns, dormant periods, missed instalments, frequency of low-balance days.	Detects changes in account behavior before formal risk events.
Transaction behavior	Cash intensity, transfer frequency, sudden spikes, unusual transaction size, inward/outward flow imbalance, activity inconsistent with profile.	Identifies deviations that may require explanation or enhanced review.
Product exposure	Number of products, credit facilities, trade finance usage, credit card exposure, overdraft limits, cross-product concentration.	Captures complexity and potential impact of customer risk.
Service interaction	Complaints, repeated limit-change requests, delayed documents, customer unreachable, unusual account-closure requests.	Adds operational and relationship signals.
Branch judgment	Site-visit concerns, adverse local information, management disputes, business closure signs, staff-coded reason flags.	Converts soft information into auditable monitoring inputs.

Statistical Classification Method

The classification method should be selected according to the bank's objective. If the goal is interpretability and quick implementation, a scorecard or logistic regression model is appropriate. If the goal is higher predictive discrimination and the bank has sufficient data, tree-based methods may be used as challenger models (Lessmann *et al*., 2015; Dumitrescu *et al*., 2022) [10, 15]. In branch banking, the recommended approach is to begin with a transparent baseline model and then test whether more complex models produce meaningful improvement without reducing usability.

A simple logistic classification model may estimate the probability that a customer will experience a defined monitoring event within a future period. The event may be a confirmed risk alert, arrears movement, account restriction, verified documentation breach, suspicious activity referral, severe complaint, or transfer to enhanced monitoring. The model can be represented as: Risk probability = f(customer profile, account conduct, transaction behavior, product exposure, service interaction and branch judgment). Customers are then grouped into risk bands using thresholds determined through historical validation and managerial capacity.

For banks with limited historical event labels, a two-stage method can be used. First, expert-defined rules classify customers into provisional monitoring bands. Second, the bank compares those bands with actual outcomes over time and recalibrates the weights. This approach is weaker than a fully validated supervised model, but it is more realistic for branches that do not yet have clean labeled data. Over time, the bank can build a reliable dataset for statistical validation.

Threshold setting is as important as model choice. If the high-risk threshold is too low, branch staff may be overwhelmed by false alerts. If it is too high, important cases may be missed. Thresholds should therefore be chosen using both statistical performance and operational workload. For example, a branch may decide that no more than a fixed percentage of customers should be escalated for enhanced review each month unless portfolio conditions deteriorate. Such a capacity constraint prevents the model from generating an unmanageable alert volume.

Class imbalance must be handled carefully. Serious customer risk events are usually rare, which makes recall, precision and capture-rate analysis more informative than overall accuracy (Brown and Mues, 2012; Chen *et al*., 2024) [7, 9]. A model can achieve high overall accuracy by classifying most customers as low risk, but such a model may fail to identify the cases that matter. Evaluation should therefore include recall for high-risk events, precision of high-risk alerts, top-decile capture rate, area under the precision-recall curve and the number of alerts per branch employee. These measures connect statistical classification to branch usefulness.

The model should also produce reason codes because explainable outputs support adoption and review in credit-risk settings (Bussmann *et al*., 2021) [8]. A reason code is a short explanation of why the customer received a risk classification. Examples include: account turnover declined by more than 40 percent, overdraft utilization remained above 90 percent for three consecutive months, KYC document expired, cash transactions increased sharply relative to profile, repeated returned cheques occurred, or branch staff recorded customer unreachable. Reason codes make the model operationally understandable.

Risk Bands and Branch Actions

A risk-based monitoring system must connect classification outputs with actions. This paper proposes three core bands and one temporary review status. The low-risk band indicates that no unusual monitoring signal is present and routine branch review is sufficient. The medium-risk band indicates that one or more early signals require verification. The high-risk band indicates persistent or multi-domain risk and requires formal escalation. The review-pending status is used when data are missing, inconsistent or awaiting branch confirmation.

Low-risk customers should not be ignored; they should continue under normal monitoring cycles. Medium-risk customers should receive proportionate attention, such as a customer call, document update, transaction explanation or review of facility usage. High-risk customers should be reviewed by the branch manager and, where appropriate, referred to the risk, credit or compliance department. Review-pending customers should not remain unresolved indefinitely; the branch should have a defined time limit to complete missing information.

The branch action plan should include ownership, time frame and evidence. A classification model will not improve governance if no one is responsible for acting on the output. Therefore, each risk band should have a named owner. Relationship managers may handle medium-risk reviews, branch managers should approve high-risk actions, and risk/compliance teams should review escalated cases. The action should be recorded in the core banking or customer relationship system so that later reviewers can see what was done.

Escalation discipline is important. Commercial pressure may lead staff to delay action against profitable customers, while excessive caution may lead to unnecessary restrictions. The classification framework should prevent both extremes by using proportionate monitoring. The objective is not to punish customers but to align branch attention with risk level and evidence.

Table 2: Illustrative classification bands and branch actions

Risk band	Classification meaning	Primary branch action	Suggested time frame
Low	No material deviation from expected profile or account behavior.	Routine monitoring; update customer information at normal review date.	Standard cycle.
Medium	One or more early signals requiring verification.	Relationship manager contacts customer, updates explanation and checks documentation.	Within 10 working days.
High	Persistent, severe or multi-domain monitoring concern.	Branch manager review; enhanced monitoring; referral to risk, credit or compliance where applicable.	Immediate to 5 working days.
Review pending	Insufficient or inconsistent data prevents reliable classification.	Resolve missing information, confirm customer profile and document decision.	Within 15 working days.

Validation and Governance

Validation is the process of checking whether the classification system works as intended. For a branch monitoring classifier, validation should consider discrimination, calibration, stability, workload, interpretability and fairness. Discrimination measures whether the model ranks riskier customers above safer customers. Calibration checks whether predicted probabilities or risk bands correspond to observed outcomes. Stability checks whether the model behaves consistently over time and across branches. Workload analysis checks whether the alert volume can be managed by available staff.

Out-of-time testing is recommended. A model developed on one historical period should be tested on a later period rather than only on randomly split data. This is important because branch behavior changes with economic conditions, digital banking adoption, policy changes and local market shocks. If the classifier performs only in the development period but fails later, it is not suitable for operational use.

Override governance is necessary because human judgment is both valuable and risky. Branch officers may know legitimate reasons for unusual behavior, but they may also override warnings due to customer pressure or relationship bias. Every override should record the direction of override, reason, evidence, approving officer and final outcome. Periodic review of overrides can reveal whether the model is too strict, whether staff are misusing discretion or whether certain branches need training.

Fairness review should also be included. A classification model may unintentionally produce different false-positive or false-negative rates across customer groups, sectors, regions or branch types. Some differences may reflect real risk variation, but others may reflect data quality or historical bias. Banks should review whether customers are being escalated for reasons that are explainable, relevant and consistent with policy.

Governance responsibility should be shared. Business banking or retail banking may own branch action, risk management may own model validation, compliance may review regulatory aspects, and internal audit may periodically test whether procedures are followed. Senior management should receive a concise monitoring report showing the number of customers by risk band, top drivers, branch workload, escalation completion and post-review outcomes.

Table 3: Statistical validation measures for a branch monitoring classifier

Validation area	Recommended measures	Managerial relevance
Discrimination	AUROC, precision-recall AUC, KS statistic, top-decile capture rate.	Shows whether riskier customers are ranked above lower-risk customers.
Calibration	Observed event rate by band, calibration plot, and Brier score.	Checks whether risk bands match actual outcomes.
Class imbalance	Recall, precision, false-negative rate, alert capture at high-risk thresholds.	Prevents misleading reliance on overall accuracy.
Workload	Number of medium/high alerts per branch employee; review completion rate.	Ensures that model outputs are operationally manageable.
Stability	Score migration, population stability index, and feature drift over time.	Detects whether the model becomes outdated.
Fairness and segmentation	Performance by branch, region, customer segment, product type and business sector.	Identifies unjustified differences and data-quality problems.

Implementation Roadmap

The framework can be implemented in phases. The first phase is diagnostic. The bank should map current monitoring practices, available data, branch workflows, common risk events and reporting gaps. This step identifies whether the required customer identifiers and historical records are reliable. The second phase is prototype development. The bank builds a simple classification score using the most reliable variables and defines provisional risk bands.

The third phase is validation. The prototype is tested against historical outcomes such as arrears, account closure, confirmed compliance alerts, complaint escalation or movement to enhanced monitoring. The bank should check performance by branch, customer segment and product type. The fourth phase is pilot implementation. A limited number of branches apply the model while continuing to record staff feedback and overrides. This phase is important because a statistically sound model may still fail if it does not fit branch workflow.

The fifth phase is policy approval and rollout. The bank approves thresholds, action rules, documentation requirements, staff responsibilities and reporting templates. Training should focus not only on how to read the score but also on how to record evidence and handle customer communication. The sixth phase is monitoring and recalibration. The bank periodically reviews model drift, alert outcomes, false positives, missed events, branch workload and fairness indicators.

Small banks can begin with a spreadsheet-based scorecard or dashboard if core banking integration is not immediately available. However, manual systems should have strict access control and version control. Larger banks should integrate the classifier into existing customer relationship management or risk systems so that alerts and actions are stored centrally. In both cases, the quality of governance is more important than the sophistication of technology.

Table 4: Implementation roadmap

Phase	Core activities	Expected output
Diagnostic	Map data, current monitoring practice, risk events, branch workflow and reporting gaps.	Data inventory and monitoring problem statement.
Prototype	Build initial scorecard or statistical classifier using reliable variables and clear labels.	Prototype risk bands and reason codes.
Validation	Test the classifier on historical outcomes and assess calibration, workload and segment performance.	Validation report and threshold recommendation.
Pilot	Apply the model in selected branches while collecting staff feedback and override	Pilot learning report and revised

	evidence.	procedures.
Rollout	Approve policy, responsibilities, training, reporting templates and escalation rules.	Operational monitoring framework.
Review	Track drift, false alerts, missed events, overrides, fairness and post-review outcomes.	Periodic recalibration and governance report.

Discussion

The proposed framework has practical implications for branch managers. First, it helps prioritize attention. Branch staff often face large numbers of customers and limited review time. Classification allows managers to focus on customers whose behavior shows meaningful change. Second, it creates consistency across branches. Two branches handling similar customers should not apply completely different monitoring standards unless local conditions justify the difference. Third, it improves documentation. A clear reason code and action record make it easier to explain why a customer was reviewed or escalated.

The framework also has implications for bank risk departments. Risk teams often design policies that are difficult for branches to operationalize. A classification-based framework creates a common language between head office and branches. It converts policy expectations into variables, thresholds, risk bands and action rules. This can improve risk reporting and reduce dependence on informal branch narratives.

For customers, the framework can support more proportionate treatment. A risk-based system does not mean that every unusual transaction or account change should lead to restrictive action. Instead, it encourages the bank to apply review intensity according to evidence. This can protect customers from arbitrary decisions while allowing the bank to act quickly when risk is credible.

There are also limitations. First, the framework depends on data quality. If branch data are incomplete or inconsistent, the classifier may produce unreliable results. Second, the model requires periodic review. A classifier that is not recalibrated can become outdated. Third, staff training is essential. Branch employees must understand that the model supports judgment but does not remove responsibility. Fourth, the paper does not present empirical testing because no confidential bank dataset was available. Future research should validate the framework using real branch-level customer data and compare model classifications with actual outcomes.

Another important limitation is that customer monitoring may involve sensitive personal and financial information. Banks must ensure that data use complies with privacy, customer confidentiality and regulatory requirements, including risk-based monitoring principles for banking activity (Financial Action Task Force, 2014) [11]. Variables should be relevant to risk and should not be used in ways that unfairly disadvantage customers. The framework should therefore be implemented with legal, compliance and data-protection review.

Conclusion

This paper developed a practical framework for risk-based customer monitoring using statistical classification in branch banking. The framework organizes branch customer information into profile, conduct, transaction, product, service and judgment domains. It then converts those variables into risk bands linked to routine monitoring, enhanced review, escalation and feedback. The main argument is that statistical classification becomes useful for branch banking only when it is connected to action, explanation and governance.

The paper contributes to management and banking practice by showing how branches can move from informal monitoring to structured, evidence-based review. It also emphasizes that classification models should be transparent, validated and manageable. Accuracy alone is insufficient. A branch monitoring classifier should be calibrated, explainable, stable, fair and consistent with operational capacity.

The proposed framework is suitable for banks with different levels of data maturity. Smaller branch networks may begin with a transparent scorecard, while larger banks may compare baseline models with machine-learning challengers. In all cases, management discipline, staff training, override documentation and periodic review are essential. Future empirical research should apply the framework to customer-month data from branch networks and test whether it improves early detection, reduces unmanaged risk and strengthens customer monitoring governance.

Declarations

Funding: The authors declare that no external funding was received for this manuscript. Replace this statement if funding was received.

Conflict of interest: The authors declare no conflict of interest. Replace this statement if any competing interest exists.

Data availability: No primary bank customer data were used in this conceptual and methodological manuscript. Future empirical validation should use anonymized and authorized bank-level data.

Ethical approval: Not applicable for this conceptual manuscript because it does not involve human participants, interviews or identifiable customer records. Ethical approval should be obtained before any future use of confidential customer-level data.

References

1. Abdou HA, Pointon J. Credit scoring, statistical techniques and evaluation criteria: A review of the literature. Intelligent Systems in Accounting, Finance and Management,2011:18(2-3):59-88.
2. Alonso Robisco A, Carbó Martínez JM. Measuring the model risk-adjusted performance of machine learning algorithms in credit default prediction. Financial Innovation,2022:8(1):Article 70.
3. Altman EI. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance,1968:23(4):589-609.
4. Bellotti T, Crook J. Support vector machines for credit scoring and discovery of significant features. Expert Systems with Applications,2009:36(2):3302-3308.

5. Berg T, Burg V, Gombović A, Puri M. On the rise of FinTechs: Credit scoring using digital footprints. The Review of Financial Studies,2020:33(7):2845-2897.
6. Bolton RJ, Hand DJ. Statistical fraud detection: A review. Statistical Science,2002:17(3):235-255.
7. Brown I, Mues C. An experimental comparison of classification algorithms for imbalanced credit scoring data sets. Expert Systems with Applications,2012:39(3):3446-3453.
8. Bussmann N, Giudici P, Marinelli D, Papenbrock J. Explainable machine learning in credit risk management. Computational Economics,2021:57(1):203-216.
9. Chen Y, Calabrese R, Martin-Barragan B. Interpretable machine learning for imbalanced credit scoring datasets. European Journal of Operational Research,2024:312(1):357-372.
10. Dumitrescu E, Hué S, Hurlin C, Tokpavi S. Machine learning for credit scoring: Improving logistic regression with non-linear decision-tree effects. European Journal of Operational Research,2022:297(3):1178-1192.
11. Financial Action Task Force. Guidance for a risk-based approach: The banking sector. FATF, Paris, 2014.
12. Fuster A, Goldsmith-Pinkham P, Ramadorai T, Walther A. Predictably unequal? The effects of machine learning on credit markets. The Journal of Finance,2022:77(1):5-47.
13. Gambacorta L, Huang Y, Qiu H, Wang J. How do machine learning and non-traditional data affect credit scoring? New evidence from a Chinese fintech firm. Journal of Financial Stability,2024:73:Article 101284.
14. Hand DJ, Henley WE. Statistical classification methods in consumer credit scoring: A review. Journal of the Royal Statistical Society: Series A,1997:160(3):523-541.
15. Lessmann S, Baesens B, Seow HV, Thomas LC. Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research. European Journal of Operational Research,2015:247(1):124-136.
16. Oskarsdóttir M, Bravo C. Multilayer network analysis for improved credit risk prediction. Omega,2021:105:Article 102520.
17. Thomas LC. A survey of credit and behavioural scoring: Forecasting financial risk of lending to consumers. International Journal of Forecasting,2000:16(2):149-172.