International Journal of Management and Commerce
www.managementjournal.in
Online ISSN: 2664-6854
Received: 23-06-2026, Accepted: 24-07-2026, Published: 10-08-2026
Volume 8, Issue 3, 2026, Page No. 21-29

Data-driven approaches to expanding SME credit access while managing portfolio quality

Md Rezwan UR Rahman[1], Mohammad Nurul Islam[2], Rasheda Yasmin[3], Mohammad Mohibul Alam[4], Zakaria Ferdous[5]

[1] Vice President, Eastern Bank PLC Unit Head, Bangladesh
[2] University of Dhaka Dhaka, Bangladesh
[3] University of Chittagong Chittagong, Bangladesh
[4] Ulster University London, United Kingdom
[5] International Islamic University Chittagong Chittagong, Bangladesh

DOI: https://doi.org/10.66856/ijmc.2026.8.3.8064

Abstract

Small and medium-sized enterprises (SMEs) remain underserved by formal credit markets even though they are central to employment, enterprise renewal and local economic resilience. Traditional bank underwriting often depends on collateral, audited financial statements and relationship history, which can exclude viable firms with thin files or irregular cash flows. This review develops an applied decision framework showing how banks and fintech-enabled lenders can widen SME credit access without weakening portfolio quality. Using a structured integrative review of recent peer-reviewed literature on fintech credit, machine learning-based credit assessment, alternative data, explainable artificial intelligence and SME financing constraints, the paper synthesizes evidence from finance, business analytics and operations research. The findings indicate that data-driven lending can improve borrower segmentation, reduce information asymmetry and support earlier risk detection, but only when predictive models are embedded in disciplined data governance, transparent decision rules, human override protocols, portfolio limits and continuous monitoring. The paper contributes the Responsible SME Credit Expansion and Portfolio Quality (R-SCEPQ) framework, which connects data inputs, risk modelling, explainability, decision policy, portfolio feedback and governance. For banks in emerging and developing markets, the framework offers a practical pathway for using digital traces, transaction histories and supply-chain signals to expand inclusion while preserving risk-adjusted returns and regulatory confidence.

Keywords: SME finance, credit access, portfolio quality, machine learning, fintech lending, credit risk management

Introduction

Access to formal credit is one of the most persistent constraints on SME growth, especially in emerging markets where business records are incomplete, collateral is scarce and lenders often lack reliable borrower histories. The challenge is not simply that banks are unwilling to lend. Rather, conventional underwriting tools frequently make the risk of many SMEs difficult to observe, quantify and monitor. A bank that lowers credit standards without improving risk information may increase non-performing loans; a bank that relies only on strict collateral and historical financial statements may protect the portfolio but deny credit to viable firms. The central managerial problem is therefore an access-quality dilemma: how can a lender expand SME credit while maintaining acceptable portfolio quality?

Data-driven lending has become an important response to this dilemma. Fintech lenders, platform-based finance providers and digitally mature banks increasingly use transaction records, e-commerce activity, payment histories, account turnover, sectoral patterns, behavioural signals and supply-chain information to supplement traditional financial statements. Empirical studies indicate that digital footprints and alternative data can reveal credit-relevant information that is not fully captured by conventional scoring variables (Berg *et al*., 2020; Jagtiani and Lemieux, 2019) [4, 22]. Recent evidence also shows that machine learning can improve default prediction and borrower ranking, although its benefits depend on data quality, stability, calibration and governance (Fuster *et al*., 2022; Bitetto *et al*., 2024; Gambacorta *et al*., 2024) [5, 6, 9].

For SMEs, the promise of data-driven credit is particularly important because many firms are neither clearly bankable nor clearly unbankable. Some are excluded because they lack formal collateral, audited accounts or long credit histories, while others are risky because their cash flows are volatile or vulnerable to sector shocks. Treating these groups as identical creates two losses: creditworthy firms remain outside formal finance and lenders miss profitable, risk-adjusted opportunities. Data-driven underwriting can potentially separate these groups more accurately, allowing lenders to approve a larger number of viable SMEs, apply risk-sensitive limits and prices, and monitor deterioration earlier.

However, the use of data and algorithms does not automatically produce responsible financial inclusion. Better prediction can also generate unequal outcomes, model opacity, unstable cut-offs, overfitting and reliance on variables that may proxy for sensitive or economically irrelevant characteristics. Fuster *et al*. (2022) [5] demonstrate that machine learning can change the distribution of credit outcomes across borrower groups, while Bussmann *et al*. (2021) and Kellner *et al*. (2022) [7, 23] show that interpretability is central to the practical use of advanced models in credit risk management. In the SME context, model governance is especially important because errors affect not only individual borrowers but also local employment, supplier networks and bank asset quality.

This paper develops an applied decision framework for lenders that wish to use data-driven approaches to expand SME credit while managing portfolio quality. It is designed for banks, microfinance institutions, credit guarantee schemes and fintech partnerships in markets where SME data are fragmented but digital traces are growing. The paper has three objectives: first, to synthesize recent literature on data-driven credit access, SME financing constraints and credit risk modelling; second, to identify the conditions under which data-driven approaches can support both inclusion and portfolio discipline; and third, to propose a practical framework that connects data architecture, model development, decision rules, portfolio monitoring and governance.

The contribution of the paper is conceptual and practical. Conceptually, it links financial inclusion and credit risk management as complementary objectives rather than competing policy goals. Practically, it translates scattered evidence from finance, fintech and analytics into an implementation-oriented framework: Responsible SME Credit Expansion and Portfolio Quality (R-SCEPQ). The framework shows that the value of data-driven lending lies not only in predictive accuracy but also in the integration of models with underwriting strategy, credit limits, human review, stress testing, early-warning dashboards and feedback loops. This emphasis aligns with the applied business and industry orientation of the journal and provides a bridge between academic evidence and lender practice.

Literature Review

The literature on SME credit access has traditionally emphasized information asymmetry, collateral deficiency, transaction costs and relationship lending. SMEs often lack the audited financial statements, managerial systems and legal collateral expected by formal lenders. In emerging markets, these constraints are intensified by informality, weak credit bureaus and uneven enforcement. Digital finance research adds a new dimension by showing that data generated through payments, e-commerce, mobile money, tax records and platform participation may reduce information frictions. Li *et al*. (2023), Guo *et al*. (2023), Yao and Yang (2022) and Zhang *et al*. (2023) [20, 30, 31] provide evidence that digital finance and fintech development can ease enterprise financing constraints, particularly for SMEs and private firms.

A second stream focuses on alternative data and fintech credit. Jagtiani and Lemieux (2019) [22] show that alternative data and machine learning can provide information beyond conventional credit scores in fintech lending. Berg *et al*. (2020) [4] demonstrate that simple digital footprints can have predictive power for default risk, while Berg, Fuster and Puri (2022) [5] synthesize the broader fintech-lending literature and distinguish the customer-interface, screening and monitoring channels through which technology changes credit provision. Cornelli *et al*. (2023) [10] document the growth of fintech and big-tech credit across countries, and Cornelli *et al*. (2024) [9] show that fintech small-business lenders can extend credit in areas less likely to be served by traditional lenders while maintaining loan-performance prediction. Gopal and Schnabl (2022) and Beaumont *et al*. (2025) [2, 17] further indicate that nonbank and fintech lenders may complement bank credit for small businesses, particularly when loan products relax collateral constraints or support asset formation. Sharma *et al*. (2024) [27] further show that fintech-led business models such as crowdfunding, platform lending, mobile payments and supply-chain finance are reshaping small business finance ecosystems.

A third literature examines machine learning and predictive analytics in credit risk. Credit scoring research shows that non-linear models can capture interactions among borrower characteristics and improve discrimination relative to traditional statistical models in many settings. Moscatelli *et al*. (2020) [25] find that machine learning models can improve corporate default forecasting when standard financial indicators are available. Wang *et al*. (2023), Xia *et al*. (2023), Belhadi *et al*. (2025) and Gu *et al*. (2024) [3, 14, 20] apply machine learning to SME and supply-chain credit risk, emphasizing the value of multi-source data, imbalance treatment and ensemble methods. Bitetto *et al*. (2024) [6] specifically investigate whether machine learning can predict small business credit risk and highlight the need to assess trust, stability and interpretability before operational adoption.

A fourth stream concerns explainability and model governance. In credit markets, predictive performance alone is insufficient because lenders must justify decisions to internal credit committees, regulators, auditors and borrowers. Bussmann *et al*. (2021) [7] propose explainable machine learning approaches for credit risk management, while Kellner *et al*. (2022) and Nazemi and Fabozzi (2024) [23, 26] show that interpretability matters for loss-given-default and recovery modelling. Danielsson *et al*. (2022) [11] warn that artificial intelligence can create systemic risk if models become correlated, poorly understood or weakly governed. These insights suggest that responsible data-driven SME lending must include auditability, documentation, stress testing and monitoring of model drift.

A final stream considers portfolio management. Expanding credit access is a portfolio decision, not merely a borrower-level scoring decision. An algorithm may identify a set of additional borrowers with acceptable predicted default probabilities, but portfolio outcomes depend on sector concentration, maturity structure, collateral coverage, pricing, guarantee arrangements, recovery processes and macroeconomic conditions. The literature on fintech lending and credit market competition shows that improved screening can affect approval rates, pricing and borrower composition (Fuster *et al*., 2019; Chu, 2024; Fuster *et al*., 2022) [5, 8, 13]. Therefore, a responsible access strategy must be evaluated through expected loss, risk-adjusted return, non-performing loan ratios, concentration limits and resilience under stress scenarios.

Taken together, the literature suggests three gaps. First, studies often emphasize either inclusion or risk prediction, but fewer studies integrate both within a decision framework for SME lenders. Second, many model-focused studies report accuracy metrics without translating them into lending policies that preserve portfolio quality. Third, studies on fintech and digital finance identify inclusion benefits but provide less guidance on the governance mechanisms required by banks and regulated lenders. The present paper addresses these gaps by developing a framework that treats data, models, policies and portfolio feedback as one integrated credit-management system.

Methodology

This paper adopts a structured integrative review and framework-development methodology. An integrative review is appropriate because the research question spans finance, business analytics, fintech, SME management and credit risk governance. The purpose is not to calculate a pooled effect size; rather, it is to synthesize evidence from recent empirical and methodological studies and translate that evidence into an applied framework for industry use. The review approach follows a transparent logic: identify relevant studies, screen them for recency and quality, extract themes related to credit access and portfolio quality, and integrate the themes into a decision framework.

The target literature covered peer-reviewed journal articles published from 2019 [13] to 2026, with emphasis on the most recent five years. Searches were designed around the terms SME finance, small business credit, fintech lending, alternative data, machine learning, explainable artificial intelligence, credit risk, portfolio quality, default prediction, supply-chain finance and digital finance. Priority was given to Q1 and Q2 journals in finance, banking, management, operations research, business analytics and information systems. Studies were included when they offered empirical evidence, model-based findings or strong conceptual implications relevant to the relationship between credit access and portfolio risk.

Studies were excluded if they were not peer reviewed, did not address credit access or credit risk, were older than 2019 [13], or focused only on consumer credit without transferable implications for SME lending. Because the study is an integrative review rather than a systematic meta-analysis, the literature set was used to build a rigorous conceptual synthesis rather than to claim exhaustive coverage of every publication in the field. The resulting evidence base combines high-level finance studies, fintech lending evidence, machine learning credit-risk research, SME-specific supply-chain finance studies and digital finance research on financing constraints.

The synthesis was conducted in three stages. First, each study was coded for its primary contribution: credit access, alternative data, predictive modelling, explainability, portfolio management or governance. Second, the studies were compared to identify points of convergence and tension. For example, several studies support the predictive value of alternative data, but others caution that model performance may produce unequal outcomes or fail under new macroeconomic conditions. Third, the evidence was translated into a practical framework that can be applied by lenders. The framework was refined around six components: data foundation, model development, decision policy, portfolio control, monitoring feedback and governance.

The methodological choice has two limitations. First, because this is a review and framework-development paper, it does not test a new model on proprietary bank data. Future empirical studies should validate the framework using loan-level SME application and repayment data. Second, the quality of any data-driven credit system depends on local institutional context, including credit bureau coverage, legal recovery processes, data protection rules and lender capabilities. The framework therefore provides a structured decision architecture rather than a universal algorithm. These limitations are addressed by specifying validation requirements, governance controls and context-specific implementation steps.

Because the study synthesizes published literature and does not involve interviews, surveys, proprietary loan-level records, identifiable borrower information or human participants, institutional ethics approval and informed consent were not required. No confidential bank data were accessed, and no copyrighted tables or figures from prior studies were reproduced.

Table 1: Inclusion and exclusion logic for the integrative review

Dimension	Inclusion rule	Exclusion rule
Period	Peer-reviewed studies published from 2019 [13] to 2026, with emphasis on 2021-2026.	Studies published before 2019 [13] unless needed for unavoidable theoretical background.
Publication quality	Studies from high-quality finance, banking, operations research, analytics, management and information systems journals, prioritizing Q1/Q2 outlets.	Non-peer-reviewed commentaries, predatory outlets, unverifiable working papers and purely promotional reports.
Topic relevance	SME finance, fintech lending, alternative data, credit risk, machine learning, explainable AI, portfolio monitoring or digital finance.	Studies unrelated to credit access, credit risk, enterprise finance or lending governance.
Evidence type	Empirical, model-based, systematic review or framework studies with clear implications for credit access or portfolio quality.	Purely technical studies with no lending or portfolio-management relevance.

Table 1 summarizes the screening logic used to keep the review focused on recent, peer-reviewed and credit-relevant evidence. This boundary was necessary because the paper aims to develop an applied decision framework rather than a broad descriptive review of every digital-finance topic.

Results and Discussion

The review indicates that data-driven approaches can expand SME credit access through five interrelated mechanisms. The first mechanism is information enrichment. Traditional underwriting relies heavily on collateral, audited accounts and credit history. Data-driven underwriting expands the information set by incorporating cash-flow data, payment histories, tax records, point-of-sale information, e-commerce performance, supply-chain relationships, utility payments and behavioural indicators. These additional signals can reveal repayment capacity in firms with thin formal credit histories. However, information enrichment creates value only when data are accurate, legally obtained, consistently updated and mapped to economically meaningful risk drivers.

The second mechanism is improved borrower segmentation. The credit access problem is partly a classification problem: lenders need to distinguish excluded but viable SMEs from genuinely high-risk applicants. Machine learning models can identify non-linear combinations of variables that traditional scorecards may miss. For example, a young SME with weak collateral but stable digital payments, diversified customers and consistent account turnover may be safer than a collateralized borrower with deteriorating cash flows.

This type of segmentation supports credit expansion because approval decisions become based on repayment evidence rather than on collateral alone. Yet segmentation should not be treated as a purely technical exercise. Lenders must test whether model segments are stable over time and whether the variables used are fair, explainable and relevant.

The third mechanism is more flexible decision policy. A model output is not a lending decision; it is an input into a decision policy. Responsible credit expansion requires multiple decision zones rather than a single accept-reject threshold. Low-risk applicants may be auto-approved subject to exposure limits, medium-risk applicants may be routed to enhanced review, and higher-risk applicants may receive smaller limits, secured products, guarantee-backed loans or advisory support rather than outright rejection. This policy design allows credit access to increase without mechanically shifting the entire portfolio toward higher risk. It also makes the model usable by credit officers who need clear rules for exception handling.

The fourth mechanism is portfolio-level control. Expanding SME credit responsibly requires linking borrower-level scores to portfolio appetite. Lenders should monitor approval rates, predicted default, expected loss, realized delinquency, sector concentration, vintage performance, migration rates and recovery outcomes. A data-driven access strategy should be judged by whether it improves the access-quality frontier: the set of lending policies that deliver higher inclusion for the same or lower risk-adjusted loss. This frontier perspective prevents managers from celebrating model accuracy while ignoring portfolio deterioration. It also helps senior credit committees compare alternative cut-offs, pricing schedules and limit structures.

The fifth mechanism is continuous learning. SME risk changes rapidly due to inflation, currency pressures, supply-chain disruptions, sector demand, tax changes and firm-level cash-flow shocks. A model trained on historical data may become less accurate when economic conditions shift. Continuous monitoring, champion-challenger testing, override analysis and feedback from recoveries are therefore necessary. Data-driven credit access should be designed as a learning system: every application, approval, rejection, delinquency and recovery outcome should improve future decision-making, provided that privacy and ethics requirements are respected.

Table 2: Evidence synthesis: mechanisms connecting data-driven lending to SME credit access and portfolio quality

Mechanism	Representative evidence	Implication for lenders
Alternative data and digital traces	Digital footprints and fintech lending data contain credit-relevant information beyond conventional score variables (Berg *et al*., 2020; Jagtiani and Lemieux, 2019; Gambacorta *et al*., 2024) [4, 9, 22].	Use alternative data to supplement, not replace, core repayment-capacity analysis.
Machine learning for risk ranking	Machine learning can improve SME and corporate default prediction when data are sufficient and validation is rigorous (Moscatelli *et al*., 2020; Wang *et al*., 2023; Bitetto *et al*., 2024) [6, 20, 25].	Measure performance using AUC, calibration, stability, recall of defaults and business impact.
Fintech and financing constraints	Digital finance and fintech development can reduce enterprise financing constraints, with stronger effects for SMEs and private firms (Li *et al*., 2023; Guo *et al*., 2023; Yao and Yang, 2022) [20, 30].	Target segments where conservative underwriting excludes viable borrowers.
Explainability and governance	Explainable AI and transparent model documentation are necessary for operational trust and regulatory accountability (Bussmann *et al*., 2021; Kellner *et al*., 2022; Nazemi and Fabozzi, 2024) [7, 23, 26].	Deploy explanations, audit trails and override monitoring before scaling automated decisions.
Portfolio feedback	Borrower-level predictions must be translated into expected loss, concentration, pricing and stress-test outcomes (Danielsson *et al*., 2022; Cornelli *et al*., 2023) [10, 11].	Judge expansion by access-quality frontier, not approval-rate growth alone.

The evidence summarized in Table 2 supports a cautious interpretation: data-driven lending can widen SME access only when additional information is converted into auditable decisions and monitored at portfolio level. This synthesis provides the basis for the integrated framework presented in
Figure 1.

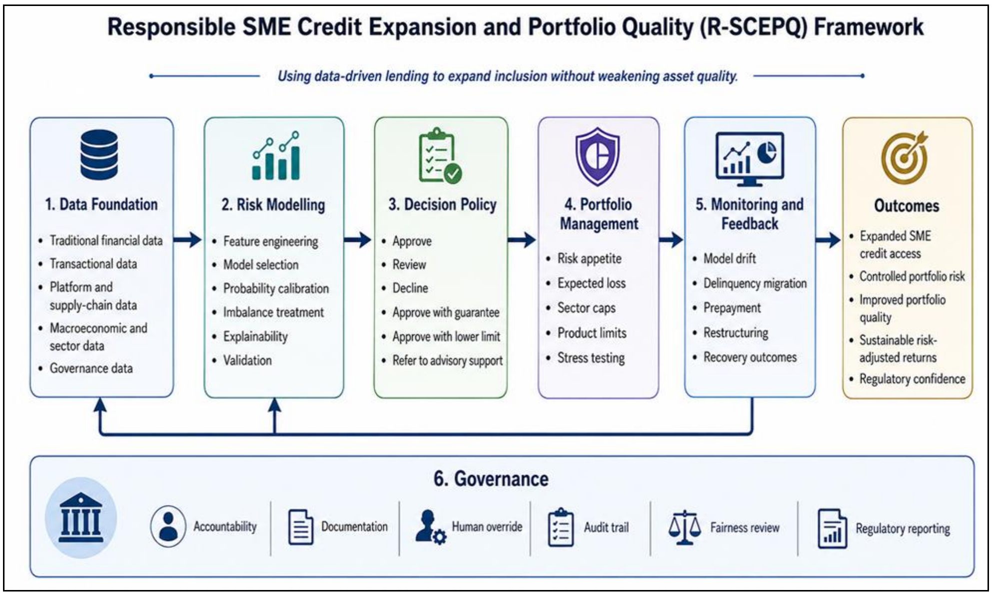

Fig 1: Responsible SME Credit Expansion and Portfolio Quality (R-SCEPQ) framework

Figure 1 should be read from left to right as a closed-loop credit operating model. Data inputs feed risk modelling, model outputs are translated into decision policy, portfolio controls constrain expansion, monitoring feeds performance information back into the system, and governance operates across every stage rather than only at approval.

Table 3: Data architecture for responsible SME credit expansion

Data layer	Examples	Control requirement
Traditional financial data	Bank statements, financial statements, tax returns, collateral records, repayment history.	Reconciliation, missing-value rules, recency checks and financial ratio validation.
Transactional data	Account turnover, cash-in/cash-out volatility, payment regularity, receivables settlement, point-of-sale flows.	Consent, source authentication, refresh frequency and outlier treatment.
Platform and supply-chain data	E-commerce ratings, invoice data, buyer concentration, delivery records, anchor-firm relationship, procurement cycle.	Data-sharing agreements, lineage controls and economic justification of variables.
Macroeconomic and sector data	Inflation, exchange rate, commodity prices, sector sales, regional demand, policy shocks.	Stress scenarios and periodic recalibration.
Governance data	Override reasons, complaint logs, rejected-application reviews, restructuring and recovery outcomes.	Audit trail, model-risk committee review and fairness checks where applicable.

Table 3 operationalizes the first component of the framework by distinguishing data layers from their control requirements. The table emphasizes that alternative data are useful only when consent, lineage, refresh frequency, economic meaning and auditability are defined before scale-up.

1. Proposed R-SCEPQ framework

The Responsible SME Credit Expansion and Portfolio Quality (R-SCEPQ) framework proposed in this paper integrates the above mechanisms into a six-part system. Component 1 is the data foundation. It defines permissible data sources, data ownership, data quality checks, refresh frequency and privacy safeguards. Component 2 is risk modelling. It covers feature engineering, model selection, probability calibration, imbalance treatment, explainability and validation. Component 3 is decision policy. It converts scores into operational decisions: approve, review, decline, approve with guarantee, approve with lower limit, or refer to advisory support. Component 4 is portfolio management. It links decisions to risk appetite, expected loss, sector caps, product limits and stress testing. Component 5 is monitoring and feedback. It captures drift, delinquency migration, prepayment, restructuring and recovery outcomes. Component 6 is governance. It establishes accountability, documentation, human override, audit trail, fairness review and regulatory reporting.

The distinctive feature of the framework is that it treats inclusion as a disciplined credit-management objective. The goal is not to maximize approval rates, nor to minimize defaults at the cost of excluding viable SMEs. The goal is to maximize sustainable access: more credit for firms with demonstrable capacity to repay, under rules that protect the lender and the financial system. This differs from simplistic digital lending models that prioritize speed without enough risk controls. It also differs from conservative bank models that protect asset quality by excluding borrowers whose risk can be measured only through non-traditional data.

The framework implies that lenders should evaluate credit expansion through combined metrics. Access metrics include approval rate, share of first-time borrowers, average processing time and geographic or sectoral coverage. Quality metrics include default rate, delinquency migration, expected loss, loss given default, recovery rate and risk-adjusted return. Governance metrics include explanation completeness, override rate, model drift, rejected-applicant review and complaint rates. A data-driven SME strategy is successful only when access improves without unacceptable deterioration in quality and governance indicators.

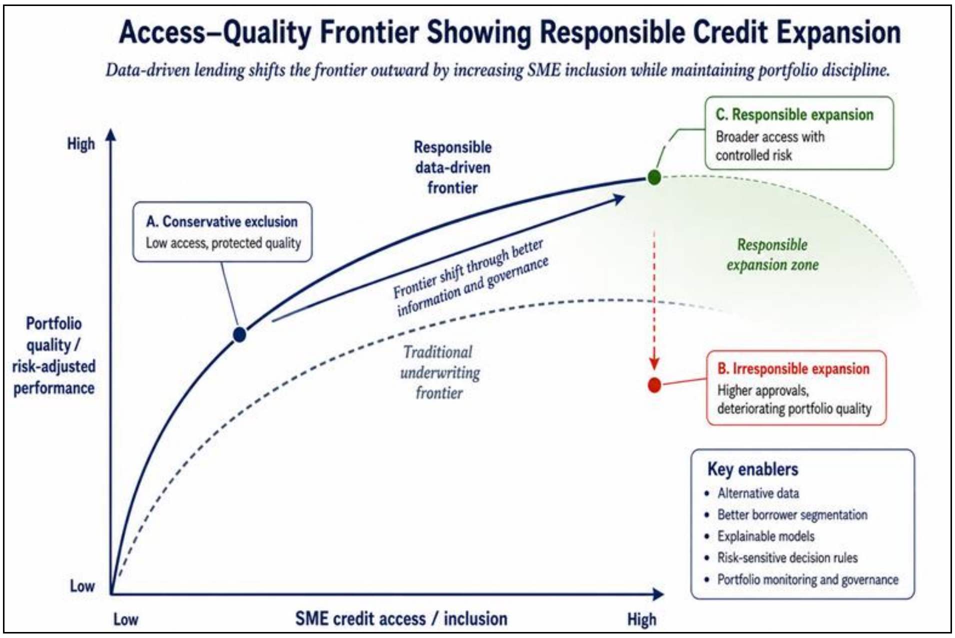

Fig 2: Access-quality frontier showing responsible credit expansion

Figure 2 clarifies the central trade-off facing lenders. A responsible strategy shifts the access-quality frontier outward by approving additional, viable SMEs without allowing expected losses, concentration risk or governance weaknesses to exceed the institution's stated risk appetite.

Table 4: Decision policy matrix for data-driven SME lending

Risk band	Indicative action	Portfolio-quality safeguard
Very low risk	Fast-track approval with standard product terms and automated documentation checks.	Exposure limits and random post-approval file audit.
Low risk	Approval with risk-sensitive pricing and normal monitoring.	Vintage tracking by sector, geography and product.
Medium risk	Enhanced review, lower initial limit, shorter tenor, partial guarantee or receivable-backed structure.	Manual confirmation of cash-flow drivers and early-warning triggers.
High but viable risk	Pilot approval only if guarantee, collateral substitute, co-lending or advisory support reduces expected loss.	Portfolio cap, monthly review and pre-defined exit criteria.
High risk	Decline or refer to non-credit support until bankability improves.	Transparent decline reason and reapplication pathway.

The decision bands in Table 4 convert model predictions into practical credit actions. They also prevent a common implementation error: treating a probability of default as a final decision rather than as one input into limits, pricing, guarantees, review intensity and borrower support.

Table 5: Portfolio monitoring dashboard for access-quality control

Metric category	Core indicators	Management use
Access	Approval rate, first-time borrower share, processing time, underserved-region coverage.	Tracks whether the strategy expands access rather than merely reprices existing borrowers.
Risk	Predicted default, realized default, 30/60/90 days past due, migration rate, expected loss.	Detects deterioration and validates model calibration.
Return	Risk-adjusted margin, fee income, recovery rate, restructuring success, cost-to-serve.	Assesses commercial sustainability.
Concentration	Sector, region, product, tenor, collateral and anchor-buyer exposures.	Prevents credit expansion from creating hidden portfolio clusters.
Governance	Override rate, explanation coverage, rejected-applicant review, model drift, complaint rate.	Ensures decisions remain auditable, fair and stable.

Table 5 closes the control loop by specifying the indicators that should be reviewed after origination. The dashboard is intentionally balanced across access, risk, return, concentration and governance so that management can detect whether expansion is genuinely responsible or merely increasing origination volume.

2. Managerial implications for banks and fintech partnerships

For banks, the framework highlights that data-driven SME lending should be implemented as a business transformation project rather than as an isolated analytics project. The first managerial task is to identify the segments where current underwriting is too conservative. These may include young firms with strong cash flows, informal firms transitioning to formalization, suppliers with stable anchor-buyer relationships, small retailers with digital payment histories and service firms with recurring receivables. The second task is to map which data sources can reduce uncertainty in those segments. The third task is to translate model outputs into credit products that fit the risk profile of each segment.

For fintech partnerships, the framework clarifies the division of responsibility. A fintech partner may provide data pipelines, scoring tools or customer acquisition channels, but the regulated lender remains responsible for risk appetite, model approval, fair treatment, data protection and portfolio outcomes. Partnership agreements should therefore include data lineage, model documentation, performance reporting, audit access, customer consent procedures and exit provisions. This is particularly important where lenders use platform data from e-commerce, ride-hailing, merchant payments or supply-chain systems. The data may be predictive, but it must be legally usable and operationally reliable.

For credit officers, the framework supports a shift from collateral-first judgment to evidence-based judgment. It does not remove human expertise; it changes the role of expertise. Credit officers should focus on interpreting exceptions, verifying unusual patterns, assessing management quality and reviewing model explanations. Human override should be permitted, but it should be documented and monitored. High override rates may indicate that the model is missing important contextual signals or that users lack trust in the model. Low override rates with rising defaults may indicate automation bias. Both cases require governance attention.

The roadmap in Table 6 is intentionally staged. Lenders with limited digital maturity can begin with diagnostic and data-readiness work, while more advanced institutions can move toward challenger models, controlled pilots and calibrated scaling once portfolio and governance thresholds are met.

Table 6: Implementation roadmap for lenders

Phase	Action	Expected output
Phase 1: Diagnostic	Identify underserved SME segments and current rejection reasons.	Baseline access-quality profile and segment opportunity map.
Phase 2: Data readiness	Audit available data sources, consent rules, data quality and integration capacity.	Approved data dictionary and governance plan.
Phase 3: Model development	Build baseline and challenger models; validate calibration, stability and explainability.	Model validation report and decision thresholds.
Phase 4: Policy design	Convert scores into approval bands, limits, pricing, guarantees and review rules.	Credit policy addendum for data-driven SME lending.
Phase 5: Pilot	Launch controlled pilot with portfolio cap and weekly monitoring.	Pilot dashboard and credit committee review.
Phase 6: Scale and learn	Scale only after performance, governance and borrower-outcome metrics are acceptable.	Ongoing monitoring, recalibration and continuous improvement.

3. Policy and ethical implications

Data-driven SME lending also has policy implications. Regulators and development institutions often seek to improve SME access through guarantees, subsidized loans and credit bureau expansion. The framework suggests that these interventions can be more effective when combined with better data infrastructure and responsible model governance. For example, a credit guarantee scheme can target medium-risk borrowers who are viable but still outside bank risk appetite. Digital tax records and e-invoicing systems can reduce information asymmetry. Open banking can allow SMEs to share transaction data with lenders, provided that consent, security and competition safeguards are strong; however, borrower-controlled data can also reshape lender competition and borrower welfare, so implementation should be accompanied by transparency and conduct safeguards (He *et al*., 2023).

Ethical safeguards are necessary because alternative data can create hidden exclusion. Variables that appear neutral may proxy for geography, gender, firm informality or platform dependence. Lenders should therefore test models for stability and disparate impact where data are available. They should avoid variables that are difficult to justify economically or legally. They should also provide meaningful explanations to declined applicants where possible. A transparent decline reason can help an SME improve future bankability, while an opaque algorithmic rejection may deepen mistrust in formal finance.

The review also suggests that responsible credit expansion must be resilient to macroeconomic shocks. Models trained during stable periods may underestimate default when inflation, exchange-rate volatility or supply-chain disruption affects SMEs. Stress testing should therefore simulate changes in sales decline, input-cost inflation, buyer concentration, interest rates and payment delays. This is especially important in emerging markets where SMEs are more exposed to macroeconomic volatility and may have limited buffers. Portfolio quality depends not only on borrower selection at origination but also on early intervention when conditions deteriorate.

4. Research implications

The framework opens several directions for future research. First, empirical studies should test whether the R-SCEPQ framework improves both access and risk outcomes using loan-level SME data. Researchers can compare incumbent scorecards with data-enriched models and evaluate approval rate, expected loss and realized performance under alternative cut-offs. Second, studies should examine which alternative data sources remain stable across sectors, regions and economic cycles. Some variables may be highly predictive in one platform or country but weak elsewhere. Third, scholars should explore how human override interacts with machine recommendations. The best outcomes may arise when algorithmic ranking is combined with disciplined human review rather than when either is used alone.

Fourth, more research is needed on fairness and inclusion in SME credit markets. Consumer credit research has advanced fairness metrics, but SME lending involves different units of analysis, such as firm size, sector, location, owner characteristics and supply-chain position. Fifth, there is a need to link credit risk modelling with product design. If a model identifies medium-risk but viable SMEs, what combination of limit, price, tenor, collateral, guarantee or advisory service best supports both repayment and business growth? These questions require collaboration between researchers, banks, fintech firms and policymakers.

Conclusion

This paper examined how data-driven approaches can expand SME credit access while managing portfolio quality. The review shows that data-driven lending can reduce information asymmetry, improve borrower segmentation, support flexible decision policy, strengthen early warning and enable continuous portfolio learning. However, the same technologies can also create opacity, unfairness, model drift and excessive risk-taking if they are implemented without governance. The core argument is that sustainable SME credit expansion requires an integrated system rather than a standalone algorithm.

The proposed R-SCEPQ framework connects six components: data foundation, risk modelling, decision policy, portfolio management, monitoring feedback and governance. This framework helps lenders move from a narrow accuracy mindset to an access-quality mindset. It encourages managers to ask not only whether a model predicts default, but also whether it enables more viable SMEs to receive appropriate credit under rules that preserve risk-adjusted portfolio performance. For emerging-market lenders, the framework is particularly relevant because digital data are expanding while traditional underwriting constraints remain significant.

The paper contributes to academic literature by linking SME financial inclusion, fintech credit, explainable machine learning and portfolio quality management. It contributes to practice by offering a structured implementation roadmap for banks and fintech-enabled lenders. Future work should validate the framework using proprietary SME loan data, compare alternative decision policies and examine how responsible data-driven lending performs during macroeconomic stress. If implemented with discipline, data-driven approaches can help lenders widen access to productive credit while protecting the resilience of the loan portfolio.

Practically, the framework should be piloted in a controlled cohort before full-scale rollout. A lender can begin by identifying one product, one SME segment and one geographic or sectoral portfolio where current rejection rates are high but observed transaction activity suggests repayable demand. The pilot should include pre-defined approval caps, manual review of borderline cases, independent model validation and weekly monitoring of early arrears, overrides and complaints. This controlled approach reduces the danger of confusing short-term origination growth with sustainable credit expansion. It also creates a feedback dataset that can be used to recalibrate scores, revise variable definitions, update cut-offs and improve explanations to credit officers and borrowers. The value of data-driven lending is therefore cumulative: it grows when each lending cycle improves bothborrower understanding and portfolio discipline.

For institutions with limited analytical infrastructure, implementation can proceed through staged maturity levels rather than immediate adoption of complex machine-learning models. At the first level, lenders can standardize data capture, define minimum documentation rules, monitor bureau and account-behavior variables, and use transparent scorecards supported by expert review. At the second level, they can incorporate digital-payment records, tax filings, supplier invoices, cash-flow seasonality and sector benchmarks, while keeping adverse-action reasons understandable to applicants. At the third level, institutions can introduce more advanced algorithms, but only after establishing model risk governance, fairness testing, audit trails, staff training and board-level risk appetite limits. This staged view is important for emerging and developing markets because many SME lenders face fragmented data, uneven borrower formality and limited specialist capacity. A responsible credit-access strategy should therefore treat data as a governance asset, not merely as a predictive input. The central managerial question is not whether a model can approve more loans, but whether it can approve additional loans in segments where repayment capacity is observable, pricing is fair, and portfolio concentration remains tolerable under stress.

Acknowledgments and Declarations

Acknowledgments: No individual or institutional acknowledgments are reported in the anonymous manuscript.

Conflict of interest: No potential conflict of interest was reported by the authors.

Funding: This work was not supported by any funding.

Authorship and contributions: All authors contributed to the conception and design of the review, literature synthesis, interpretation, manuscript drafting, critical revision and final approval of the version submitted for publication. All authors agree to be accountable for the accuracy and integrity of the work.

Responsible use of artificial intelligence: OpenAI ChatGPT was used to support manuscript structuring, drafting assistance and language refinement during the development of this manuscript. The authors take full responsibility for reviewing, verifying and editing all AI-assisted content, confirming originality, checking all citations and approving the final submitted version.

Research ethics: This integrative review did not involve human participants, identifiable personal data, clinical data, proprietary loan files or intervention with individuals. Institutional ethics approval and informed consent were therefore not required.

Data and materials availability: No new empirical datasetwas generated or analyzed for this integrative review. The literature sources used to support the synthesis are listed in the reference section

References

1. Arner DW, Buckley RP, Zetzsche DA, Veidt R. Sustainability, FinTech and financial inclusion. European Business Organization Law Review,2020:21:7-35.
2. Beaumont P, Tang H, Vansteenberghe E. Collateral effects: The role of FinTech in small business lending. The Review of Financial Studies, 2025.
3. Belhadi A, Kamble SS, Mani V, Benkhati I, Touriki FE. An ensemble machine learning approach for forecasting credit risk of agricultural SMEs' investments in agriculture 4.0 through supply chain finance. Annals of Operations Research,2025:345(2):779-807.
4. Berg T, Burg V, Gombovic A, Puri M. On the rise of FinTechs: credit scoring using digital footprints. The Review of Financial Studies,2020:33(7):2845-2897.
5. Berg T, Fuster A, Puri M. FinTech lending. Annual Review of Financial Economics,2022:14:187-207.
6. Bitetto A, Cerchiello P, Filomeni S, Tanda A, Tarantino B. Can we trust machine learning to predict the credit risk of small businesses? Review of Quantitative Finance and Accounting,2024:63(3):925-954.
7. Bussmann N, Giudici P, Marinelli D, Papenbrock J. Explainable machine learning in credit risk management. Computational Economics,2021:57(1):203-216.
8. Chu Y. Fintech lending and credit market competition. Journal of Financial and Quantitative Analysis,2024:59(5):2199-2225.
9. Cornelli G, Frost J, Gambacorta L, Jagtiani J. The impact of fintech lending on credit access for U.S. small businesses. Journal of Financial Stability,2024:73:101290.
10. Cornelli G, Frost J, Gambacorta L, Rau PR, Wardrop R, Ziegler T. Fintech and big tech credit: drivers of the growth of digital lending. Journal of Banking & Finance,2023:148:106742.
11. Danielsson J, Macrae R, Uthemann A. Artificial intelligence and systemic risk. Journal of Banking & Finance,2022:140:106290.
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. Fuster A, Plosser M, Schnabl P, Vickery J. The role of technology in mortgage lending. The Review of Financial Studies,2019:32(5):1854-1899.
14. Gambacorta L, Huang Y, Qiu H, Wang J. How do machinelearning and non-traditional data affect credit

scoring? New evidence from a Chinese fintech firm. Journal of Financial Stability,2024:73:101284.
15. Gao H, Kou G, Peng Y, Alsaadi FE. Machine learning in business and finance: a literature review and research opportunities. Financial Innovation, 202,10:86.
16. Goodell JW, Kumar S, Lim WM, Pattnaik D. Artificial intelligence and machine learning in finance: identifying foundations, themes, and research clusters from bibliometric analysis. Journal of Behavioral and Experimental Finance,2021:32:100577.
17. Gopal M, Schnabl P. The rise of finance companies and FinTech lenders in small business lending. The Review of Financial Studies,2022:35(11):4859-4901.
18. Gu F, Wu Y, Hu X, Liu Y. The impact of digital inclusive finance on SMEs' technological innovation activities: empirical evidence from China. PLOS ONE,2023:18(10):0293500.
19. Gu Z, Lv J, Wu B, Hu Z, Yu X. Credit risk assessment of small and micro enterprise based on machine learning. Heliyon,2024:10(5):27096.
20. Guo J, Fang H, Liu X, Wang C. FinTech and financing constraints of enterprises: evidence from China. Journal of International Financial Markets, Institutions and Money,2023:82:101713.
21. He Z, Huang J, Zhou J. Open banking: Credit market competition when borrowers own the data. Journal of Financial Economics,2023:147(2):449-474.
22. Jagtiani J, Lemieux C. The roles of alternative data and machine learning in fintech lending: evidence from the LendingClub consumer platform. Financial Management,2019:48(4):1009-1029.
23. Kellner R, Nagl M, Rosch D. Opening the black box - quantile neural networks for loss given default prediction. Journal of Banking & Finance,2022:134:106334.
24. Li C, Wang Y, Zhou Z. Digital finance and enterprise financing constraints: structural characteristics and mechanism identification. Journal of Business Research,2023:165:114074.
25. Moscatelli M, Parlapiano F, Narizzano S, Viggiano G. Corporate default forecasting with machine learning. Expert Systems with Applications,2020:161:113567.
26. Nazemi A, Fabozzi FJ. Interpretable machine learning for creditor recovery rates. Journal of Banking & Finance,2024:164:107187.
27. Sharma SK, Goyal DP, Sharma A. Small businesses and FinTech: a systematic review and future directions. Electronic Commerce Research,2024:24(1).
28. Wang L, Jia F, Chen L, Xu Q. Forecasting SMEs' credit risk in supply chain finance with a sampling strategy based on machine learning techniques. Annals of Operations Research,2023:331(1):1-33.
29. Xia Y, Xu T, Wei MX, Wei ZK, Tang LJ. Predicting chain's manufacturing SME credit risk in supply chain finance based on machine learning methods. Sustainability,2023:15(2):1087.
30. Yao L, Yang X. Can digital finance boost SME innovation by easing financing constraints? Evidence from Chinese GEM-listed companies. PLOS ONE,2022:17(3):0264647.
31. Zhang L, Zhang Y, Chen X. Digital inclusive finance, financing constraints, and technological innovation of SMEs. Sustainability,2023:15(9):7144.