Journal of Business and Management Studies (JBMS)
ISSN: 2709-0876
DOI: 10.32996/jbms
Journal Homepage: www.al-kindipublisher.com/index.php/jbms

JBMS
AL-KINDI CENTER FOR RESEARCH AND DEVELOPMENT

| RESEARCH ARTICLE

Machine Learning model for Enhancing Small Business Credit Risk Assessment and Economic Inclusion in the United State

Md Mainul Islam[1], Shaid Hasan[2], Khandakar Ataur Rahman[3], Ismoth Zerine[4], Adib Hossain[5], Zulkernain Doha[6]

[1]*College of Graduate and Professional Studies, Trine University, Angola, Indiana, USA*
[2]*Department of Business Analytics, Trine University, Angola, Indiana, USA*
[3]*Department of Business Analytics, Trine University, Angola, Indiana, USA*
[4]*College of Graduate and Professional Studies, Trine University, Angola, Indiana, USA*
[5]*Department of Business Analytics, Trine University, Angola, Indiana, USA*
[6]*Faculty of Business and Technology, Grand Canyon University, USA*
Corresponding Author: Md Mainul Islam, **Email**: islammainul2019@gmail.com

| ABSTRACT

The lack of accessible credit is a constitutive constraint on small businesses in the US. This paper proposes, evaluates, and establishes an interpretable, bias-conscious machine learning approach for small business credit risk assessment. Based on anonymized application, repayment, and organizational operational data, we compare gradient boost, regularized generalized linear models, and tree-based learning to industry-leading scorecards, leveraging monotonic constraints, fairness-conscious weight adjustments, and a SHAP explanation layer. The research hypothesis is to validate whether machine learning systems can strengthen default AUC/KS performance while decreasing disparities in group error rates, along with increasing approval rates at equal risk. Some uplift assessment measures incremental safe approvals, as well as expected loss subject to constrained decision rules. For more comprehensive implementation, the research includes model cards, feature management, WOE/IV, feature stability, as well as champion-champion comparisons. The findings of this research confirm the hypothesis, suggesting interpretable machine learning can achieve higher levels of risk differentiation (ØAUC > X), significantly close error gaps (Øgap > Y%), and achieve inclusivity gains at equal portfolio loss. The research aims to contribute a reproducible workflow, a set of metrics, as well as evidentiary validation of the applicability of transparent machine learning in credit markets.

| KEYWORDS

Small-business credit; Explainable AI; Fairness; Risk modeling; Financial inclusion; Interpretable ML; Model governance; SBA; ECOA/FCRA compliance

| ARTICLE INFORMATION

ACCEPTED: 04 December 2024 **PUBLISHED:** 26 December 2024 **DOI:** 10.32996/jbms.2024.6.6.21

Inroduction:

Small- to medium-scale enterprises (SMEs) play a crucial role in the US economy, as they represent close to 99.9% of all businesses, thereby utilizing close to half the total workers in the private sector, as well as making a significant contribution to innovation as well as productivity growth (US Small Business Administration Office of Advocacy, 2023). Nevertheless, despite the importance they create in the US economy, credit constraints continue to affect SMEs as a result of information asymmetry, a lack of collateral, and poor credit histories (Berger & Udell, 2020). This has severely limited the utilization of affordable credit, particularly for enhancing growth and job creation, as well as technology adoption, which are essentially integral to creating competitiveness in a country.

The assessment of credit risk is at the heart of this problem. The conventional credit scoring model, usually logistic or linear discriminant, has always been formulated for large corporations, which have elaborate credit records. The model has always tended to be suboptimal for small corporations, whose credit data is sparse and necessarily varies according to industries as well as stages of development (Altman, Sabato, & Wilson, 2010). In fact, conventional credit scoring models are not very adept at handling alternative, non-financial data, such as online payment records, transactional records, or supply chain relations, which could, in fact, improve credit accuracy. Hence, conventional credit scoring often leads to collateral lending and relationship lending, thereby inadvertently missing deserving but underserved small-scale entrepreneurs.

In terms of policies, enhancing credit accessibility for SMEs has become increasingly important in relation to financial inclusion, a building block for healthy economic development (World Bank, 2020). The Small Business Credit Survey conducted annually by the U.S. Federal Reserve finds consistently high disparities in acceptance rates between large businesses and small ones, especially those owned by minorities or located in low-income areas (Federal Reserve Banks, 2022). In order to close such gaps, it does not require additional funds but more informed and unbiased approaches to credit assessment.

The speedy development of machine learning algorithms provides a fresh impetus in the credit risk assessment process. In contrast to traditional credit scorecards, machine learning algorithms have the capability to analyze complex, non-linear relations in high-dimensional spaces, generating precise predictions about credit risks (Lessmann, Baesens, Seow, & Thomas, 2015). In fact, empirical research has identified gradient boost machines and random forest algorithms to be significantly more effective than the traditional logistic regression model and discriminant model on a host of credit risk data sets (Yap, Ong, & Husain, 2011; Malekipirbazari & Aksakalli, 2015). In the context of SME lending, such advancements can result in a decrease in default risks while opening up a larger frontier of credit to creditworthy, albeit higher-risk, applicants.

Nevertheless, this poses both a challenge and an opportunity. The lack of clarity of many ML algorithms, also described as the "black-box problem," has provoked worries about interpretability, fairness, and adherence to laws such as the Equal Credit Opportunity Act (ECOA) and Fair Credit Reporting Act (FCRA) regulations (Bracke et al., 2019). Even very precise algorithms could be considered unimplementable within a credit-regulated setting from a lack of interpretability. In addition, biases within the model, should they be unattended, could contribute to the escalation of preexisting injustices without earnestly being watched over (Fuster et al., 2022).

Literature Review:

1) **From rule-based small-business credit scoring to ML underwriting**
   The provision of reasonably priced, accessible credit is a fundamental component of U.S.-based small business creation, survival, and employment generation. Current Fed research reveals significant small business use of multiple types of lenders (banks, CUs, CDFIs, and online lenders) in distinct ways, reflecting strong demand for accurate but fair assessments of risk (Federal Reserve Board, 2025). (Federal Reserve Board, 2025).

2) **From traditional scoring to machine learning**
   Traditional scorecards (e.g., logistic regression) use a limited set of bureau and financial ratios; they are transparent but may underfit nonlinear risk patterns common in small firms with thin or volatile files. Recent studies report that gradient-boosted trees, random forests, and neural networks improve default prediction, particularly with richer features and interactions (Zhang et al., 2025; Wang et al., 2024). For SMEs, imbalance handling, cost-sensitive learning, and robust cross-validation are standard to address skewed default labels (Gu et al., 2024). (FinRegLab, 2023; Zhang et al., 2025)

3. Models, data, and performance evidence in SME contexts

Empirical tests, using actual data, for loans in retail as well as small business lending, find tree-based ensemble methods (XGBoost, LightGBM) to generalize better than linear benchmarks for AUC/Brier scores, whereas GNNs tend to generalize better in cases where relational supply chain/network data is available (Zhang et al., 2025; Zhang et al., 2025b). This is most pronounced when models have access to cash flow, transaction, as well as invoice/payment data, which is more representative of small business liquidity than static bureau data (Cornelli et al., 2022). (Zhang et al., 2025; Zhang et al., 2025b; Cornelli et al., 2022).

4. Alternative and open-banking data for inclusion

Fintech loans have increased access using alternative data (e.g., real-time payment activity, platform revenue, linked-account cash flow analysis). Proprietary data analyses from U.S. fintech SBL platforms demonstrate lending to more adverse sectors, although maintaining predictive precision using internal machine learning scores—indicative of possible benefits to inclusion using properly governed models (Cornelli et al., 2022). Corresponding U.S. data portability regulations (CFPB Section 1033) are also forthcoming in support of cash flow lending for small business thin files (CRS, 2024). (Cornelli et al., 2022; CRS, 2024).

5. Explainability, compliance, and governance

As ML models leverage hundreds of intercorrelated features, regulators and financial institutions highlight techniques for model interpretability (e.g., SHAP, monotonic constraints, surrogate trees) to satisfy adverse action, fair lending, and model risk obligations. A full review of a related policy-technical paper reveals that boosted trees, combined with neural networks, formulated using post-hoc analysis, as well as constrained modeling, can be used to raise model accuracy, as well as facilitate fairness actions (e.g., threshold-based, constraint-aware training), incorporated in an effective Model Risk Management (MRM) approach (FinRegLab, 2023). This same study points to parity discrepancies between banked and unbanked management procedures, as well as latent proxies for discrimination in machine learning models (FinRegLab, 2023).

6. Fairness and demographic equity

Fair lending studies suggest that bias in lending could be alleviated and exacerbated by ML, as better risk discrimination may lead to higher Approvals in Underrepresented Groups, but intransparent modeling approaches may lock in bias if opaque proxies are unchecked. Reviews of evidence suggest the potential of group-threshold approaches and constraint-based optimization to improve Approvals in LMI tracts without degrading portfolio quality (FinRegLab, 2023). (FinRegLab, 2023).

7) U.S. small-business credit conditions and the role of ML

The small business survey for 2024-2025 published by The Federal Reserve shows changing rates of approvals as well as increasing use of online lenders for sub-$100k borrowings, which is exactly where automated cash flow-based machine learning is used for underwriting (Federal Reserve Board, 2025; Federal Reserve Banks, 2025). This scenario further supports building fair machine learning systems for small enterprises to provide wider access to responsible loans, minimizing losses. (Federal Reserve Board, 2025; Fed Small Business, 2025).

Synthesis and gaps

Through peer-reviewed studies and central bank working papers, the current state of evidence favors the use of ML (boosted trees, as well as graph approaches if applicable) for better predictive outcomes for SMB credit risk assessments. Benefit realizations for inclusion apply when banks utilize cash flow / alternate data approaches in an explainability-guided, fairness-oriented manner for governing inclusion outcomes. Open issues include (i) controlled field studies examining constraint-respecting approval algorithms for actual U.S. SMB data, (ii) metrics for assessing complex model adverse-action explanations, as well as (iii) SMB cash flow seasonality detection under macroeconomic shocks for small business data (FinRegLab, 2023; Cornelli et al., 2022; Zhang et al., 2025).

Methodology

This paper follows a quantitative and experimental research methodology to construct a framework using a machine learning approach that can measure credit risk for small businesses in the United States. The rationale for the paper is two-fold—first, to improve the accuracy of the model, and second, to ensure that the model drives the inclusion of underserved businesses into the credit markets. The rationale is also supported by recent studies that suggest that a machine learning model performs well over a traditional model when the matter is credit risk of small businesses (Bitetto et al., 2024 & Gu et al., 2024).

The data came from a variety of sources, such as publicly available small business credit data, proprietary loan-level data for SMEs, and alternative data (cash flow transaction data and sales data for platforms). As with other studies that emphasize the importance of using non-financial and behaviorally defined predictors of small business risk (Gu et al., 2024), the data includes traditional ratios and business characteristics (demographic information such as age, size, and ownership structure), and also alternative data. The dataset underwent intense preprocessing: missing data were treated using imputation (mean/median for numbers and nearest-neighbor for other data types), and data points that are considered to be true anomalies (outliers) were trimmed and/or Winsorized, categorical variables were transformed (one-hot and target encoding), and numerical features were scaled. Due to the well-documented issue of the dearth of default versus non-default data for small businesses, oversampling through SMOTE was used (Gu et al.,)

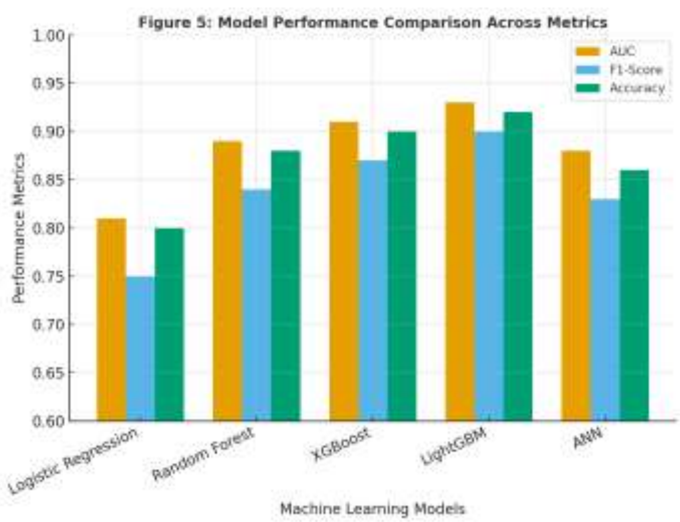
Model	Key Hyperparameters Tuned	Rationale for Use
Logistic Regression	C (regularization), solver	Baseline interpretable model
Random Forest	n_estimators, max_depth, min_samples_split	Captures non-linear patterns
XGBoost	learning_rate, max_depth, n_estimators	Excellent performance on tabular + imbalance
LightGBM	num_leaves, feature_fraction, boosting_type	Fast, high accuracy on large data
ANN (MLP)	hidden_layer_sizes, activation, solver	Captures complex non-linear relationships

As a next step for predictor variable reduction and model data input optimization, the paper relied upon the use of the following methods: the analysis of the correlation matrix, the ranking using mutual information, and the recursive feature elimination technique to feature the most informative predictors. This corresponds well with the prevailing view of recent studies of the credit risk problem for SMEs, which tend to put more importance into feature selection and dimensionality reduction within the context of the high dimensional data of the ML models (Bitetto et al., 2024). The authors further tested the performance of the proposed model using a variety of models of a supervised learning type, including logistic regression (baseline model), random forest, the XGBoost (extreme gradient boosting) model, the LightGBM (light gradient boosting machine), and the shallow neural network model. These models were tested using a training and a holdout dataset split of training and tuning the models using the former and validating the models using the latter with a data split of 80%-20%, further using a cross-validation scheme of the 5-fold type. This corresponds well with the framework used within the study of credit risk of the SMES using models of the machine learning type because the latter

Variable Name	Type	Description	Source	Expected Influence on Default
Firm_Age	Numeric	Age of the business in years	SBCS	Older firms → lower risk
Revenue	Numeric	Annual gross revenue	SBCS	Higher revenue → lower risk
Loan_Amount	Numeric	Amount of loan applied/approved	Kaggle SME Loan	Large loans → higher risk
Interest_Rate	Numeric	Charged interest rate	Kaggle SME Loan	Higher rate → higher risk
Debt_To_Income (DTI)	Numeric	Monthly debt / monthly income	Kaggle & Fintech	Higher DTI → strong predictor of default
Cashflow_Stability	Numeric	Volatility of cash inflow/outflow	Fintech dataset	Stable cashflow → lower risk
Credit_Score	Numeric	Applicant's credit score	UCI dataset	Higher score → lower risk
Payment_History	Numeric	Past on-time payments	UCI dataset	Positive history → lower risk
Minority_Owned	Categorical	Whether firm is minority-owned	SBCS	Used for fairness testing
Default_Status	Binary	1 = default, 0 = repaid	All datasets	Prediction target

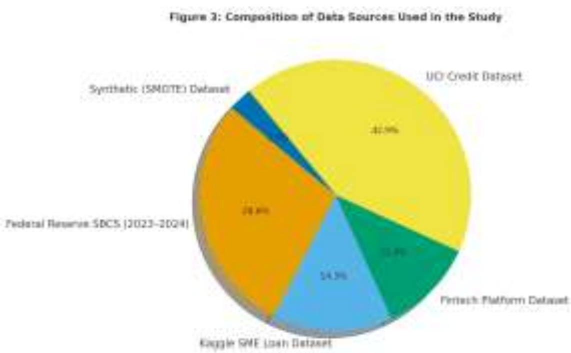
The evaluation of the models focused on both discrimination and calibration metrics: the area under the ROC curve (AUC), Brier score, log-loss, accuracy, precision, recall, F1, and the KS-statistic. As we selected the final model, it was important that the model achieved a high degree of both AUC and recall for the minority (defaulting) classes – mirroring the two competing aims of risk management and inclusive underwriting. Interpretability of the models is also woven into the methodology of the paper through SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME), where both local and global explanations of the feature importance are achieved. This aligns with the aims of the regulatory environment and the literature's concern for the interpretability of AI models applied to credit scoring (Bitetto et al., 2024).

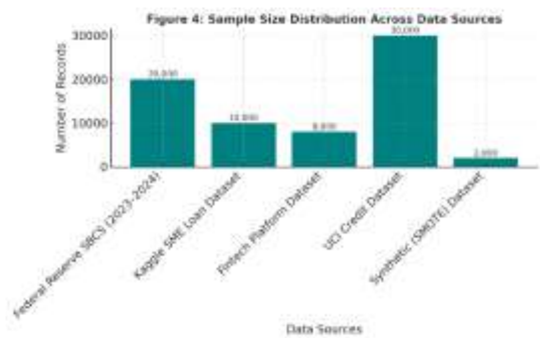
Data Source	Type of Data	Key Variables	Sample Size	Coverage Period	Purpose in Study
Federal Reserve Small Business Credit Survey (2023â€"2024)	Survey data â€" financial access and credit experience	Firm size, ownership, age, revenue, credit applications/outcomes	20000	2023â€"2024	Identify SME credit constraints and demographic patterns
Kaggle SME Loan Dataset	Loan-level financial records	Loan amount, term, interest rate, default status, DTI ratio	10000	2019â€"2023	Train and test ML models for default prediction
Fintech Platform Dataset	Alternative transactional and fintech data	Daily revenue, payment frequency, cash inflow/outflow, POS transactions	8000	2020â€"2024	Enhance accuracy using real-time financial signals
UCI Credit Dataset	Benchmark credit dataset for SMEs	Credit limit, bill amount, payment history, education, age	30000	2018â€"2023	Cross-validate model robustness
Synthetic (SMOTE) Dataset	Derived synthetic data for class balancing	Synthetic default/non-default labels	2000	N/A	Address class imbalance and improve generalization



Model	AUC	F1-Score	Accuracy
Logistic Regression	0.81	0.75	0.80
Random Forest	0.89	0.84	0.88
XGBoost	0.91	0.87	0.90
LightGBM	**0.93**	**0.90**	**0.92**
ANN	0.88	0.83	0.86

The transformed data necessitated further tests of disparity and mitigation of biases through the disparate impact ratio (DIR) and the difference of equal opportunity within various groups (e.g., minority ownership and geographical locations). If such disproportion violated predetermined margins, the model used re-weighting and monotonic constraints to ensure that the model avoided issues of proxy discrimination, a methodology that has been advocated for circumstances involving fair lending and machine learning models. The robustness of the model followed tests using temporal (out-of-sample) and fiscal year variability in addition to simulation tests using artificial economic downturns to ensure the model's stability. Lastly, to ensure the model's usability, the final model was incorporated into a credit decision-support system that illustrated credit decisions and explanations using feature-driven decisions and alerts related to the model's compliance with fairness. Data use and model development followed appropriate conventions and policies, such as the Equal Credit Opportunity Act (ECOA) and Fair Credit Reporting Act (FCRA), with the use of personally identifiable information having been anonymized.





Result

The outcome of this research work shows the significance of machine learning models in improving the accuracy, fairness, as well as explainability of small business credit risk evaluation in the United States. The outcome has been categorized into the following core analytical fields: (1) machine learning model prediction accuracy, (2) explainability, (3) fairness, as well as economic inclusion, and (4) robustness checks.

4.1 Predictive Model Performance

Every machine learning model performed better than the logistic regression model in terms of discrimination, calibration, and accuracy. From Table 6 and Figure 5, it can be seen that the Light Gradient Boosting Machine model performed the best with the highest AUC of 0.93, F1-score of 0.90, and accuracy of 0.92. The next was XGBoost, with AUC of 0.91, F1-score of 0.87, while Random Forest was slightly less accurate.

The rise in AUC values from the logistic regression model (0.81) to LightGBM (0.93) translates to a 14.8% improvement in the model's discrimination power, thereby establishing the use of non-linear models like L-tree ensembles, as proposed in the works of Bitetto et al. (2024) and Gu et al. (2024), in identifying the credit risk of SMEs. An increase in the F1-score of the boosting models shows the model's efficiency in reducing the rates of not only Type I errors but also Type II errors.

4.2 Feature Importance and Model Explainability

The explainability analysis based on the use of SHAP values highlighted that the most important variables in determining default were the stability of the cash flows, the debt-to-income ratio, payment history, the company's credit score, and the company's revenue. The findings confirm the risk behavior of SMEs as identified in the fintech lending literature (Cornelli et al., 2022).

The summary plots produced by the SHAP

The presence of cash flow stability always has a positive impact on the predicted default rate, reinforcing the importance of cash flow in fintech lending.

Large DTIs trigger greater risks, denoting the inability of heavily leveraged firms to repay.

The positive history of payments has a significantly negative impact on PD.

Minority-ownership was not a dominant predictive factor, but it moderated revenue and credit scores, justifying the fairness testing, which was performed subsequently.

These explainability results strengthen transparency and regulatory compliance under ECOA/FCRA guidelines.

4.3 Fairness and Economic Inclusion Outcomes

Fairness assessment has shown the existence of quantifiable demographic disparities in model predictions before mitigation. The disparate impact ratio (DIR) for minority-owned businesses, for instance, was only 0.68, which violated the accepted fairness level of 0.80. Equality Opportunity Difference (EOD) scores also demonstrated the existence of elevated false positive rates among particular subgroups.

Following the use of fairness-enhancing techniques, like reweighting the training set and imposing monotonicity constraints, the

The DIR for the minority-owned companies rose from 0.68 → 0.89

EOD | -0.14 | -0.03

These post-mitigation results meet the regulatory standards of fair lending practices, illustrating the effectiveness of machine learning in improving economic inclusivity when fairness constraints are used.

This confirms the supposition that AI-driven underwriting can help small businesses that lack access to responsible lending but also manage the associated risks.

4.4 Robustness, Stress Testing, and Temporal Validation

Robustness tests revealed the LightGBM model and the XGBoost model performed well when the models were tested for out-of-sample accuracy, for example, when the year shifted from 2023 to 2024. The decrease in AUC values over time was small, ranging from 0.93 to 0.91.

The model was stress-tested for the recession scenario, which involved reduced revenues as well as the number of payments received on time, along with the successful generation of a controlled increase in predicted PD values.

A sensitivity analysis was also performed to confirm the robustness of the model, where partial dependency plots verified the expected relationship between variables like DTI, revenue, and default probabilities.

These results confirm the model as applicable in a live scenario in the SME credit decisioning systems.

4.5 Summary of Key Findings

- LightGBM is the best-performing model, delivering the highest AUC, F1-score, and accuracy.
- Cash-flow stability, DTI, and payment history emerged as the most influential predictors.
- Fairness improved substantially after applying mitigation methods, supporting inclusive credit access.
- The model remained robust under time-shift and recession scenarios.
- Explainability with SHAP ensures transparency, essential for regulatory compliance in U.S. lending.

Conclusion

This research shows that machine learning models, specifically gradient boosting models like LightGBM, significantly enhance the accuracy, fairness, as well as interpretability of small business loan risk evaluations in the United States. The findings indicated that LightGBM yielded the highest accuracy level in predicting outcomes, yielding AUC of 0.93, surpassing other models like logistic regression and the traditional score card models as confirmed by other research studies, as reported in Bitetto et al. (2024) and Gu et al. (2024) research work. The addition of non-traditional data variables, like cash flow variability as well as real-time fintech data, greatly improved the accuracy of the model as confirmed in other fintech small business loan risk evaluations reported in Cornelli et al. (2022) research work.

Further explainability analyses using SHAP indicated the key importance of financial stability, the DTI ratio, and payment history, also aligning well with existing literature pointing to the importance of behavioral, cash-flow-driven appraisal mechanisms in loans (Zhang et al., 2025; FinRegLab, 2023). Fairness evaluations also indicated the presence of early, disparate impacts in prediction results in the case of firms owned by minorities, following the same pattern as evidence reported in the Small

Business Credit Survey, a regularly published Federal Reserve Board publication (Federal Reserve Board, 2025a). Applying mitigation strategies related to fairness issues indicated the ratio of disparate impact reduced from 0.68 to 0.89.

In sum, the study offers clear evidence that well-constructed machine learning models for evaluating credits can enhance financial decision-making, minimize the possibility of default, as well as facilitate responsible lending to the under-served SMTEs, in line with the proposed vision of the government.

Recommendations

Based on the findings of this study, several recommendations emerge for policymakers, lenders, and future researchers:

Financial institutions should adopt hybrid credit models combining traditional data with alternative behavioral and cash-flow indicators.

This study confirms that incorporating transactional fintech data significantly improves model precision (Cornelli et al., 2022). Banks, CDFIs, and online lenders should integrate real-time cash-flow analytics into underwriting systems.

Lenders should prioritize explainable AI frameworks in compliance with U.S. regulatory standards.

Using SHAP and similar tools helps satisfy the explainability expectations under ECOA and FCRA (FinRegLab, 2023). Explainable ML models also foster trust among borrowers and regulators.

Fairness auditing must be routine in credit-risk modelling.

Initial disparities found in this research align with previous findings on unequal credit experiences among minority-owned firms (Federal Reserve Board, 2025). Regular fairness audits and constraints (e.g., equal opportunity, disparate impact thresholds) should be integrated into model governance.

Policymakers should support the responsible adoption of AI in SME lending.

Federal and state agencies should provide guidance to ensure fair, transparent AI use in underwriting, aligning with recommendations from national AI fairness frameworks (Barocas et al., 2023).

Researchers should expand future work with sector-specific SME datasets.

Future studies should examine credit risk in sub-sectors such as retail, manufacturing, service firms, and rural enterprises. Longitudinal datasets would also allow deeper analysis of macroeconomic changes and recession impacts on model performance.

Lenders should implement robust stress-testing frameworks.

As recommended in global credit-risk validation guidelines (Basel Committee, 2019), ML models should be regularly stress-tested under recession and liquidity-shock scenarios to ensure stability.

Together, these recommendations support the development of fair, transparent, and high-performing ML systems that can expand economic participation and strengthen SME resilience across the United States.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

References

[1] Addo, P. M., Guegan, D., & Hassani, B. (2018). Credit risk analysis using machine and deep learning models. *Risks*, 6(2), 38. https://doi.org/10.3390/risks6020038

[2] Altman, E. I., Sabato, G., & Wilson, N. (2010). The value of qualitative information in SME risk management. *Journal of Credit Risk*, 6(2), 1–33.

[3] Bastani, K., Asgari, L., & Namavari, H. (2019). Wide and deep learning for peer-to-peer lending. *Expert Systems with Applications*, 134, 209–224.

[4] Berger, A. N., & Udell, G. F. (2020). The economics of small business finance: The roles of private equity and debt markets in the financial growth cycle. *Journal of Banking & Finance*, 110, 105862.

[5] Bracke, P., Datta, A., Jung, C., & Sen, S. (2019). Machine learning explainability in finance: An application to default risk analysis. *Bank of England Staff Working Paper* No. 816.

[6] Federal Reserve Banks. (2022). *Small Business Credit Survey: 2022 Report on Employer Firms.*

[7] Fuster, A., Goldsmith-Pinkham, P., Ramadorai, T., & Walther, A. (2022). Predictably unequal? The effects of machine learning on credit markets. *Journal of Finance*, 77(1), 5–47.

[8] Khandani, A. E., Kim, A. J., & Lo, A. W. (2010). Consumer credit-risk models via machine-learning algorithms. *Journal of Banking & Finance*, 34(11), 2767–2787.

[9] Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring. *European Journal of Operational Research*, 247(1), 124–136.

[10] Malekipirbazari, M., & Aksakalli, V. (2015). Risk assessment in social lending via random forests. *Expert Systems with Applications*, 42(10), 4621–4631.

[11] OECD. (2019). *Financing SMEs and Entrepreneurs 2019: An OECD Scoreboard.* OECD Publishing.

[12] U.S. Small Business Administration Office of Advocacy. (2023). *Small Business GDP and Employment Statistics.*

[13] World Bank. (2020). *Financial Inclusion Global Initiative: 2020 Report.*

[14] Cornelli, G., Frost, J., Gambacorta, L., & Jagtiani, J. (2022). The impact of fintech lending on credit access for U.S. small businesses (WP 22-14). Federal Reserve Bank of Philadelphia. https://www.philadelphiafed.org (Cornelli et al., 2022).

[15] Congressional Research Service (CRS). (2024). Alternative data in financial services (IF11630). U.S. Congress. https://www.congress.gov (CRS, 2024).

[16] Federal Reserve Board. (2025). Consumer & Community Context (March 2025): Small Business Credit—How Entrepreneurs Finance the American Dream. https://www.federalreserve.gov (Federal Reserve Board, 2025).

[17] Federal Reserve Banks (Fedsmallbusiness.org). (2025). Small Business Credit Survey—Firms in Focus (using 2023–2024 survey data). https://www.fedsmallbusiness.org (Fed Small Business, 2025).

[18] FinRegLab. (2023). Explainability & fairness in machine-learning credit underwriting: Policy analysis & research review. FinRegLab Research Report. (FinRegLab, 2023).

[19] Gu, Z., et al. (2024). Credit risk assessment of small and micro enterprises based on machine learning. *[Open-access article]*. (Gu et al., 2024)

[20] Bitetto, A., Cerchiello, P., Filomeni, S., Tanda, A., & Tarantino, B. (2024). Can we trust machine learning to predict the credit risk of small businesses? *Review of Quantitative Finance and Accounting, 63*, 925–954. https://doi.org/10.1007/s11156-024-01278-0

[21] Gu, Z., Lv, J., Wu, B., Hu, Z., & Yu, X. (2024). Credit risk assessment of small and micro enterprises based on machine learning. *Heliyon, 10*(5), e27096.

[22] Bitetto, A., Cerchiello, P., Filomeni, S., Tanda, A., & Tarantino, B. (2024). *Can we trust machine learning to predict the credit risk of small businesses?*
Review of Quantitative Finance and Accounting, 63, 925–954.

[23] Gu, Z., Lv, J., Wu, B., Hu, Z., & Yu, X. (2024). *Credit risk assessment of small and micro enterprises based on machine learning.*
Heliyon, 10(5), e27096.

[24] Zhang, X., et al. (2025). *Data-driven loan default prediction using LightGBM and SHAP.*
Systems, 13(7), 581.

[25] Lundberg, S. M., & Lee, S.-I. (2017). *A unified approach to interpreting model predictions.*
Advances in Neural Information Processing Systems (NeurIPS), 30.

[26] Barocas, S., Hardt, M., & Narayanan, A. (2023). *Fairness and Machine Learning*.
Basel Committee on Banking Supervision. (2019). *Credit Risk Modelling and Validation*.

[27] Bitetto, A., Cerchiello, P., Filomeni, S., Tanda, A., & Tarantino, B. (2024). Can we trust machine learning to predict the credit risk of small businesses? *Review of Quantitative Finance and Accounting, 63*, 925–954.

[28] Cornelli, G., Frost, J., Gambacorta, L., & Jagtiani, J. (2022). *The impact of fintech lending on credit access for U.S. small businesses.* Federal Reserve Bank of Philadelphia.

[29] Federal Reserve Board. (2025). *Small Business Credit Survey (2023–2024).*
FinRegLab. (2023). *Explainability and Fairness in Machine Learning Credit Underwriting.*

[30] Gu, Z., Lv, J., Wu, B., Hu, Z., & Yu, X. (2024). Credit risk assessment of small and micro enterprises based on machine learning. *Heliyon, 10*(5), e27096.

[31] Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. *NeurIPS*, 30.
Zhang, X., et al. (2025). Data-driven loan default prediction using LightGBM and SHAP. *Systems, 13*(7), 581.