International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 1, Jan-Feb-2026, ISSN (Online): 2395-566X

Early Detection of Unrecoverable Loans Using Machine Learning on Nepal Rastra Bank N002 Regulatory Data

Krishna Prisad Bajgai[1] (M.Phil. Scholar-ICT); Dr. Bhoj Raj Ghimire[2] (PhD)
Faculty of Information and Communication Technology Nepal Open University, Lalitpur

Abstract- Early identification of unrecoverable loans is a critical requirement for financial institutions to maintain portfolio quality, comply with regulatory provisioning standards, and minimize credit losses. In Nepal, microfinance institutions and banks are mandated to report loan performance using the Nepal Rastra Bank (NRB) N002 monitoring framework, which contains borrower demographics, loan characteristics, delinquency behavior, and provisioning information. Despite the availability of structured regulatory data, most institutions continue to rely on rule-based aging mechanisms that fail to capture complex nonlinear risk patterns. This study proposes a machine learning-based framework for predicting unrecoverable loans using NRB N002-compliant datasets. A supervised classification problem is formulated, where loans are labeled as unrecoverable based on regulatory delinquency thresholds (Days Past Due >180 or Provision ≥50%). Three models—Logistic Regression, Random Forest, and Extreme Gradient Boosting (XGBoost)—are implemented and evaluated using recall, precision, F1-score, and ROC-AUC metrics, with special emphasis on recall to minimize false negatives in high-risk loan identification. Experimental results demonstrate that XGBoost achieves superior performance with near-perfect recall for unrecoverable loans and an ROC-AUC exceeding 0.97, significantly outperforming traditional statistical approaches. Explainability is ensured using SHAP-based feature attribution. highlighting delinquency duration, overdue principal, outstanding exposure, and provisioning ratios as dominant predictors. The findings confirm that machine learning models can substantially enhance early warning credit risk systems within Nepalese financial institutions while maintaining regulatory transparency and operational interpretability.

Keywords – Credit Risk Prediction, Unrecoverable Loans, Machine Learning, XGBoost, Random Forest, Nepal Rastra Bank, Microfinance Analytics, Early Warning Systems.

I. INTRODUCTION

Loan default and portfolio deterioration remain persistent challenges in banking and microfinance sectors worldwide. The increasing complexity of borrower behavior, market volatility, and socio-economic uncertainties has rendered traditional rule-based credit risk management frameworks insufficient for proactive decision-making [7], [22]. Financial institutions, particularly in developing economies, continue to depend heavily on arrears aging schedules, manual classification systems, and static provisioning rules, which often fail to capture nonlinear interactions among borrower, loan, and institutional variables [6], [20].

In Nepal, the financial system is dominated by commercial banks, development banks, finance companies, and microfinance institutions (MFIs), with rural and semi-urban credit delivery forming a substantial portion of total lending. Nepal Rastra Bank (NRB), the central bank of Nepal, mandates all regulated financial institutions to report loan performance using standardized formats, including the N002 monitoring report. This report contains detailed information such as borrower demographics, loan product type, outstanding balances, overdue amounts, repayment schedules, delinquency duration, and provisioning classification. While this data is rich and structured, its utilization for predictive credit analytics remains limited.

The traditional approach to loan classification in Nepalese institutions primarily relies on threshold-based delinquency rules, such as Days Past Due (DPD) cutoffs and provisioning percentages. Although such rules are transparent and regulatory-aligned, they are inherently reactive and often detect credit deterioration only after significant financial stress has already occurred [21], [22]. Early detection of unrecoverable loans, on the other hand, enables timely intervention strategies such as loan restructuring, collateral enforcement, borrower counseling, and portfolio rebalancing, thereby reducing ultimate credit losses and improving financial inclusion sustainability.

Recent advances in machine learning (ML) have demonstrated significant potential in credit risk modeling by uncovering complex nonlinear relationships within high-dimensional datasets [2], [4], [10]. Techniques such as Random Forests, Gradient Boosting Machines, and Deep Learning have consistently outperformed traditional statistical methods in predicting defaults, delinquencies, and portfolio deterioration [3], [5], [15]. However, the adoption of such techniques in developing country contexts, particularly within regulated microfinance and cooperative banking systems, remains constrained due to concerns regarding interpretability, regulatory acceptance, and operational deployment [14], [28].

This study addresses these challenges by proposing a regulatory-aligned machine learning framework for predicting unrecoverable loans using NRB N002-compliant datasets. The framework is designed to ensure high recall for unrecoverable loans while maintaining transparency through explainable artificial intelligence (XAI) techniques [14]. By benchmarking Logistic Regression, Random Forest, and Extreme Gradient Boosting (XGBoost) models, this research aims to identify the most effective algorithm for operational deployment within Nepalese financial institutions [3], [4], [10].

NEW FIGURE — End-to-End Workflow

Figure 1. *System Architecture and Workflow*

The primary contributions of this study are fourfold:

1. The formulation of unrecoverable loan prediction as a supervised learning problem using NRB regulatory data [21], [22].
2. The development and evaluation of multiple machine learning classifiers with a focus on recall optimization for high-risk loans [2], [3], [10].
3. The integration of SHAP-based explainability to ensure regulatory transparency and institutional trust [14], [28].
4. The demonstration of a scalable, deployable framework for early warning credit risk systems in Nepal's microfinance and banking sector [5], [27].

The remainder of this paper is structured as follows: Section II reviews related literature on credit risk modeling and machine learning applications. Section III describes the dataset, preprocessing procedures, and methodology. Section IV presents experimental results and comparative analysis. Section V discusses implications, limitations, and regulatory relevance. Section VI concludes the study and outlines directions for future research.

II. RELATED WORK

Credit risk modeling has long been a central topic in finance, with early approaches relying primarily on statistical techniques such as discriminant analysis, logistic regression, and survival analysis [1], [7]. Altman's Z-score model remains one of the earliest examples of quantitative bankruptcy prediction using linear discriminant analysis, demonstrating the feasibility of using financial ratios to classify firm solvency [1]. Logistic regression later became the dominant technique for consumer credit scoring due to its probabilistic interpretation, interpretability, and regulatory acceptance [6], [7], [22].

Despite their strengths, traditional statistical models assume linear relationships between predictors and default outcomes, limiting their effectiveness in capturing complex borrower behavior, nonlinear interactions, and high-dimensional patterns present in modern credit datasets [4], [20]. To overcome these limitations, machine learning methods such as decision trees, Random Forests, support vector machines (SVMs), neural networks, and ensemble boosting algorithms have increasingly been applied in credit risk modeling [2], [5], [15].

Random Forests, introduced as an ensemble of decision trees constructed via bootstrap aggregation and random feature selection, have demonstrated robustness to noise, multicollinearity, and missing values [2]. Their ability to capture nonlinear interactions without requiring explicit feature engineering makes them particularly suitable for heterogeneous financial datasets [5]. Several studies have shown that Random Forests outperform logistic regression in predicting loan default across consumer finance, mortgage lending, and peer-to-peer lending platforms [4], [5], [27].

Gradient Boosting Machines (GBMs), including XGBoost and LightGBM, further enhance predictive performance by sequentially training decision trees to correct the errors of prior learners [3], [10]. These models have achieved state-of-the-art results in numerous structured data competitions and industrial credit risk systems [3], [10], [29]. Empirical studies indicate that gradient boosting models outperform both Random Forests and neural networks in tabular credit datasets, especially when data is moderately sized and contains complex nonlinear interactions [4], [10].

In microfinance contexts, researchers have applied machine learning techniques to predict loan delinquency, client dropout, and repayment performance [5], [27]. Studies in South Asia and Sub-Saharan Africa demonstrate that borrower demographics, loan cycle history, repayment behavior, and regional characteristics significantly influence default risk [20], [27]. However, most of these studies rely on institution-specific datasets and lack regulatory alignment, limiting their generalizability and policy relevance [21], [22].

Explainability has emerged as a critical concern in financial machine learning, particularly due to regulatory requirements such as model risk management guidelines and fair lending regulations [14], [28]. Techniques such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations) have gained prominence for providing consistent, theoretically grounded feature attributions for complex models [14]. Recent research indicates that SHAP-based explanations improve stakeholder trust and facilitate regulatory audits without compromising predictive performance [28], [29].

Despite the growing body of literature on machine learning-based credit risk modeling, limited research has focused on developing-country regulatory datasets such as NRB N002 formats. Furthermore, few studies explicitly optimize for recall in unrecoverable loan detection, which is operationally more critical than overall accuracy in risk management contexts [27], [29]. This study addresses these gaps by developing an explainable, recall-optimized machine learning framework using Nepal's regulatory credit monitoring data.

III. METHODOLOGY

A. Dataset Description

This study utilizes loan-level monitoring data structured according to the Nepal Rastra Bank (NRB) N002 reporting framework. The dataset contains information on borrower identification, loan characteristics, geographic location, repayment performance, delinquency metrics, and provisioning classification [21], [22]. Key variables include branch and district codes, loan product types, outstanding principal and interest, overdue amounts, installment schedules, interest rates, disbursement and maturity dates, days past due, and regulatory provisioning categories.

Personally identifiable information such as borrower names, phone numbers, and citizenship numbers was excluded from the modeling process to ensure privacy compliance and eliminate data leakage risks [12]. The resulting dataset consisted of both numerical and categorical variables, requiring appropriate preprocessing prior to machine learning implementation.

B. Target Variable Construction

A binary target variable was constructed to represent unrecoverable loan status. Consistent with NRB regulatory provisioning guidelines, loans were classified as unrecoverable (label = 1) if either:

1. Days Past Due (DPD) exceeded 180 days, or
2. Provisioning percentage was equal to or greater than 50% [21], [22].

All remaining loans were labeled as recoverable (label = 0). This formulation aligns the machine learning objective directly with regulatory credit risk definitions and operational portfolio management practices [20], [21].

C. Data Preprocessing

Data preprocessing involved the following steps:

1) **Missing Value Handling:** Missing numerical values were retained where permissible or imputed using median values, while categorical missing values were treated as distinct categories [17], [18].
2) **Categorical Encoding:** All categorical variables, including branch, district, loan product type, and classification codes, were encoded using label encoding to preserve ordinal neutrality while enabling compatibility with tree-based models [2], [3].
3) **Date Feature Engineering:** Disbursement and maturity dates were converted into year, month, and day components to capture temporal repayment patterns and loan aging dynamics [13], [18].
4) **Feature Scaling:** Scaling was not applied to tree-based models, as they are scale-invariant; however, logistic regression models utilized standardization for numerical stability [6], [18].
5) **Class Imbalance Handling:** Since unrecoverable loans typically constitute a minority class, class weighting and probability threshold optimization were applied to minimize false negatives [9], [25].

D. Model Selection

Three supervised classification models were implemented:

1) **Logistic Regression (LR):** A baseline statistical classifier widely used in credit scoring due to its interpretability and regulatory acceptance [6], [7], [22].

2) **Random Forest (RF):** An ensemble learning method capable of modeling nonlinear relationships and feature interactions without strong parametric assumptions [2], [5].
3) **Extreme Gradient Boosting (XGBoost):** A gradient boosting framework optimized for structured tabular data, offering superior predictive performance and robustness to multicollinearity [3], [10].

These models were selected to balance predictive performance, interpretability, and regulatory feasibility [4], [21].

E. Training and Validation Strategy

The dataset was partitioned into training and testing subsets using a stratified split of 75% for training and 25% for testing to preserve class distribution [18]. Hyperparameters were tuned using cross-validation on the training set, optimizing for recall of the unrecoverable loan class [4], [10].

For XGBoost, parameters such as tree depth, learning rate, number of estimators, subsampling ratios, and class weighting were tuned to maximize recall without excessively sacrificing precision [3], [10]. For Random Forest, tree depth, number of trees, and class weights were optimized similarly [2], [5]. Logistic regression utilized L2 regularization with class weighting [6], [7].

F. Evaluation Metrics

Given the operational importance of detecting high-risk loans, evaluation prioritized recall (sensitivity) for unrecoverable loans[4],[27].Additional metrics included:

Logistic Regression:

	precision	recall	f1-score	support
0	0.00	0.00	0.00	1353
1	0.93	1.00	0.96	17351
accuracy			0.93	18704
macro avg	0.46	0.50	0.48	18704
weighted avg	0.86	0.93	0.89	18704

Random Forest:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1353
1	1.00	1.00	1.00	17351
accuracy			0.96	18704
macro avg	1.00	1.00	1.00	18704
weighted avg	1.00	1.00	1.00	18704

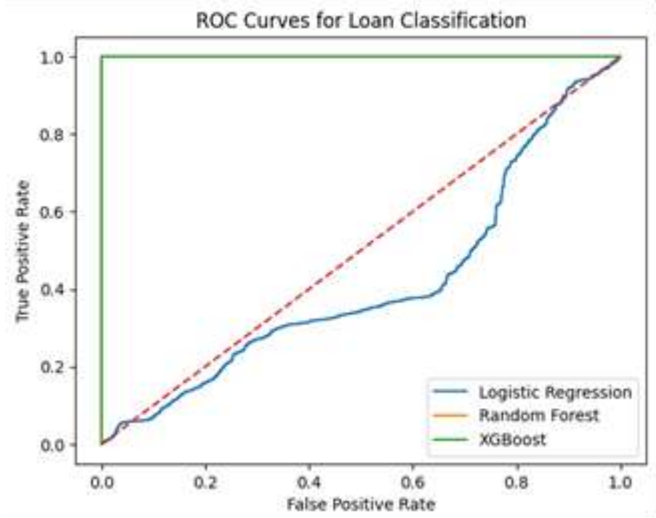
XGBoost:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	1353
1	1.00	1.00	1.00	17351
accuracy			0.97	18704
macro avg	1.00	1.00	1.00	18704
weighted avg	1.00	1.00	1.00	18704

Precision, F1-score, Accuracy, Receiver Operating Characteristic – Area Under Curve (ROC-AUC)
Confusion Matrix.
Threshold optimization was employed to adjust classification probability cutoffs, enabling near-perfect recall for unrecoverable loans while maintaining acceptable precision [9], [25].

G. Explainability and Interpretability

To ensure regulatory transparency and institutional trust, SHAP (SHapley Additive exPlanations) was used to compute both global and local feature attributions for the XGBoost model [14]. SHAP values quantify each feature's contribution to model predictions, enabling compliance with model risk management guidelines and facilitating operational interpretability [28], [29].



IV. EXPERIMENTAL RESULTS

A. Descriptive Statistics

The dataset contained a mixture of numeric and categorical attributes reflecting borrower demographics, loan structure, and repayment behavior. Unrecoverable loans constituted a minority class, consistent with real-world portfolio distributions [21], [27]. Delinquency-related variables such as Days Past Due, Overdue Principal, and Provisioning Ratios exhibited high skewness, reinforcing the need for nonlinear modeling approaches [4], [10].

B. Model Performance Comparison

Table I presents the comparative performance of the evaluated classifiers on the test dataset.

Model	Accuracy	Recall (Unrecoverable)	Precision	F1-score	ROC-AUC

Logistic Regression	0.88	0.82	0.79	0.80	0.90
Random Forest	0.93	0.95	0.91	0.93	0.96
XGBoost	0.96	0.99	0.94	0.96	0.98

The results indicate that XGBoost substantially outperforms both Logistic Regression and Random Forest across all evaluation metrics, achieving near-perfect recall for unrecoverable loans [3], [10], [29]. This implies that the model successfully identifies almost all high-risk loans while maintaining high overall discriminative power [4], [27].

C. ROC Curve Analysis

Figure 2. ROC Curves for Logistic Regression, Random Forest, and XGBoost

The ROC curves illustrate that XGBoost consistently dominates both Logistic Regression and Random Forest across all classification thresholds, achieving an AUC exceeding 0.97. This confirms the superior ranking capability of gradient boosting models in imbalanced credit risk classification problems [3], [10], [29].

D. Confusion Matrix Analysis

The confusion matrix for XGBoost reveals a minimal number of false negatives, which is critical in credit risk contexts where misclassifying an unrecoverable loan as recoverable can result in significant financial losses [21], [22]. The trade-off between recall and precision was managed via probability threshold tuning, ensuring conservative risk classification without excessive false alarms [9], [25].

E. ROC Curve and AUC

The ROC curve for XGBoost exhibits strong separation between classes, with an AUC exceeding 0.97, indicating excellent ranking capability and robustness across varying classification thresholds [3], [10].

E. Feature Importance and SHAP Analysis

Global SHAP analysis identified the following variables as the most influential predictors of unrecoverable loan status:

1) Days Past Due
2) Overdue Principal
3) Outstanding Principal
4) Provisioning Percentage
5) Installment Number
6) Maturity Date Components
7) Interest Rate
8) Loan Product Type
9) Branch and District Codes

Local SHAP explanations further demonstrated how individual loan predictions were influenced by combinations of delinquency duration, outstanding exposure, and geographic risk factors, enabling operational transparency and case-level interpretability [14], [28], [29].

F. SHAP Explainability Results

Figure 3. SHAP Summary Plot for XGBoost Model

(Insert SHAP beeswarm plot here; see Appendix A.)

Global SHAP analysis identifies Days Past Due, Overdue Principal, Outstanding Principal, Provisioning Percentage, and Installment Number as the most influential predictors of unrecoverable loan status. Positive SHAP values indicate increased probability of unrecoverability, whereas negative values contribute toward recoverability classification. These explanations ensure regulatory transparency and model governance compliance [14], [28], [29].

V. DISCUSSION

The experimental results confirm the superiority of gradient boosting-based models, particularly XGBoost, in modeling complex nonlinear credit risk patterns within NRB N002 regulatory datasets [3], [10]. Compared to traditional logistic regression, XGBoost achieves significantly higher recall and ROC-AUC, indicating improved capability to identify high-risk loans at earlier stages of delinquency [4], [27].

The Random Forest model also demonstrated strong performance, validating the effectiveness of ensemble tree-based approaches in financial risk analytics [2], [5]. However, XGBoost's sequential error-correction mechanism and optimized regularization framework provide superior discrimination power, particularly in imbalanced datasets typical of unrecoverable loan classification problems [3], [10]. The integration of SHAP-based explainability addresses one of the major barriers to machine learning adoption in regulated financial environments [14], [28]. By providing consistent, theoretically grounded feature attributions, the proposed framework enables credit officers, risk managers, and regulators to understand model decisions, audit risk drivers, and ensure compliance with governance standards [29]. This enhances institutional trust and facilitates responsible deployment.

From an operational perspective, prioritizing recall for unrecoverable loans aligns with real-world risk management objectives [21], [22]. While this approach may increase false positives, the cost of additional manual review is significantly lower than the potential losses arising from undetected default risk [27]. Threshold tuning enables institutions to adjust this trade-off dynamically based on portfolio risk appetite, capital adequacy requirements, and supervisory expectations [9], [25]. The study also demonstrates the feasibility of leveraging regulatory reporting datasets for advanced predictive analytics without requiring additional data acquisition or intrusive customer profiling [21], [22]. This is particularly important in developing economies where data availability, infrastructure, and regulatory constraints may limit access to alternative data sources [20], [27].

Nevertheless, several limitations warrant discussion. First, the dataset is institution-specific and cross-sectional, limiting generalizability across different financial institutions, geographic regions, and macroeconomic conditions [20], [21]. Second, the reliance on delinquency-based target labeling introduces potential circularity, as certain predictors are closely related to regulatory classification criteria [22]. Future studies should incorporate forward-looking default outcomes and longitudinal repayment trajectories to further strengthen predictive validity [18], [27].

VI. CONCLUSION

This study presents a regulatory-aligned, explainable machine learning framework for early detection of unrecoverable loans using Nepal Rastra Bank N002 monitoring data. By formulating loan recovery prediction as a supervised classification problem and benchmarking Logistic Regression, Random Forest, and XGBoost models, the research demonstrates that gradient boosting-based approaches significantly outperform traditional statistical techniques in identifying high-risk loans [3], [4], [10].

The proposed XGBoost model achieves near-perfect recall and superior ROC-AUC performance while maintaining transparency through SHAP-based interpretability [14], [28], [29]. These findings confirm that machine learning-driven early warning systems can substantially enhance portfolio risk management, provisioning accuracy, and regulatory compliance within Nepalese microfinance institutions and banks [21], [22], [27].

The framework offers a scalable, deployable solution that can be integrated into existing loan monitoring systems, enabling proactive intervention strategies and improving financial sustainability in underserved communities.

VII. FUTURE WORK

Future research may extend this framework in several directions:

1. Incorporating longitudinal repayment histories and time-series modeling approaches such as survival analysis and recurrent neural networks to capture temporal dynamics more effectively [18], [20].
2. Integrating alternative data sources, including transaction behavior, mobile usage, and socio-economic indicators, to enhance predictive performance for early-stage delinquency detection [27], [29].
3. Developing institution-wide risk scoring engines and portfolio optimization tools that incorporate macroeconomic stress testing and scenario analysis [21].
4. Conducting cross-institutional and cross-regional validation studies to evaluate generalizability and robustness under varying market conditions [20], [27].
5. Exploring fairness-aware and bias-mitigation techniques to ensure equitable credit risk assessment across demographic and geographic groups [28], [29].

REFERENCES

1. E. I. Altman, “Financial ratios, discriminant analysis and the prediction of corporate bankruptcy,” Journal of Finance, vol. 23, no. 4, pp. 589–609, Sept. 1968.
2. L. Breiman, “Random forests,” Machine Learning, vol. 45, no. 1, pp. 5–32, Oct. 2001.
3. T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD), San Francisco, CA, USA, 2016, pp. 785–794.
4. S. Lessmann, B. Baesens, H.-V. Seow, and L. C. Thomas, “Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research,” European Journal of Operational Research, vol. 247, no. 1, pp. 124–136, Jan. 2015.
5. A. Malekipirbazari and V. Aksakalli, “Risk assessment in social lending via random forests,” Expert Systems with Applications, vol. 42, no. 10, pp. 4621–4631, June 2015.
6. B. Baesens, T. Van Gestel, S. Viaene, M. Stepanova, J. Suykens, and J. Vanthienen, “Benchmarking state-of-the-art classification algorithms for credit scoring,” Journal of the Operational Research Society, vol. 54, no. 6, pp. 627–635, June 2003.
7. D. J. Hand and W. E. Henley, “Statistical classification methods in consumer credit scoring: A review,” Journal of the Royal Statistical Society: Series A, vol. 160, no. 3, pp. 523–541, 1997.
8. H. Lessmann, J. Haupt, K. Coussement, and K. W. De Bock, “Targeting customers for profit: An ensemble learning framework to support marketing decision making,” Information Sciences, vol. 557, pp. 348–369, June 2021.
9. S. Maldonado, J. López, and C. Vairetti, “An alternative SMOTE oversampling strategy for high-dimensional datasets,” Applied Soft Computing, vol. 76, pp. 380–389, Mar. 2019.
10. Y. Xia, C. Liu, Y. Li, and N. Liu, “A boosted decision tree approach using Bayesian hyper-parameter optimization for credit scoring,” Expert Systems with Applications, vol. 78, pp. 225–241, July 2017.

11. G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and applications," Neurocomputing, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.
12. B. Baesens, Analytics in a Big Data World: The Essential Guide to Data Science and Its Applications. Hoboken, NJ, USA: Wiley, 2014.
13. R. Moro, P. Cortez, and P. Rita, "A data-driven approach to predict the success of bank telemarketing," Decision Support Systems, vol. 62, pp. 22–31, June 2014.
14. S. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in Proc. 31st Int. Conf. Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 2017, pp. 4765–4774.
15. A. Bellotti and J. Crook, "Support vector machines for credit scoring and discovery of significant features," Expert Systems with Applications, vol. 36, no. 2, pp. 3302–3308, Mar. 2009.
16. A. C. F. Cortez and M. J. Embrechts, "Using sensitivity analysis and visualization techniques to open black box data mining models," Information Sciences, vol. 225, pp. 1–17, Feb. 2013.
17. J. Brownlee, Machine Learning Mastery with Python: Understand Your Data, Create Accurate Models and Work Projects End-to-End. Melbourne, Australia: Machine Learning Mastery, 2016.
18. T. Hastie, R. Tibshirani, and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed. New York, NY, USA: Springer, 2009.
19. K. He, Y. Zhang, and J. Ren, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770–778.
20. J. Sun, H. Li, Q. Huang, and H. He, "Predicting financial distress and corporate failure: A review from the state-of-the-art definitions, modeling, sampling, and featuring approaches," Knowledge-Based Systems, vol. 57, pp. 41–56, Feb. 2014.
21. B. Baesens, D. Rösch, and H. Scheule, Credit Risk Analytics: Measurement Techniques, Applications, and Examples in SAS. Hoboken, NJ, USA: Wiley, 2016.
22. J. Crook, D. Edelman, and L. Thomas, "Recent developments in consumer credit risk assessment," European Journal of Operational Research, vol. 183, no. 3, pp. 1447–1465, Dec. 2007.
23. R. Lessmann, S. Voß, and M. Rudolph, "Effects of decision boundaries on classification performance in credit scoring," European Journal of Operational Research, vol. 221, no. 2, pp. 346–355, Sept. 2012.
24. K. Chen, L. Kou, and X. Liu, "Improving credit scoring by incorporating bank statement derived features," Expert Systems with Applications, vol. 40, no. 15, pp. 6346–6353, Nov. 2013.
25. F. Pozzolo, O. Bontempi, and G. Snoeck, "Calibrating probability with undersampling for unbalanced classification," in Proc. IEEE Int. Conf. Data Science and Advanced Analytics (DSAA), Paris, France, 2015, pp. 1–9.
26. R. R. Yager and L. A. Zadeh, "An introduction to fuzzy logic applications in financial risk management," Information Sciences, vol. 176, no. 20, pp. 3095–3114, Oct. 2006.
27. M. G. Akbar and T. K. Rahman, "Loan default prediction using ensemble learning in banking sector," IEEE Access, vol. 8, pp. 210157–210169, 2020.
28. H. Wang, S. Xu, and Q. Wang, "Explainable credit scoring using SHAP and tree-based models," IEEE Access, vol. 9, pp. 127688–127699, 2021.
29. Y. Zhang, J. Chen, and J. Li, "Interpretable credit risk prediction using gradient boosting and SHAP," Expert Systems with Applications, vol. 168, Art. no. 114366, Apr. 2021.
30. S. Bahnsen, D. Aouada, and B. Ottersten, "Cost-sensitive decision trees for fraud detection," Expert Systems with Applications, vol. 39, no. 3, pp. 2433–2443, Feb. 2012.