

American Journal of Scholarly Research and Innovation

Volume: *4;* ***Issue:*** *1*
Pages: *1174– 1202*
eISSN: 3067-2163

Machine Learning for Nonperforming Corporate Loan Recovery: A Systematic Review and Meta-Analysis of Predictive Performance

Md Hasinur Rahman[1]

[1]. Independent Banking & Financial Risk Researcher;
Master of Arts in Professional Communication, University of San Francisco, California, USA;
Email: rahmanhasinur10@gmail.com

Doi: 10.63125/a7v6f533
Received: *15 September 2025;* ***Revised:*** *21 October 2025;* ***Accepted:*** *10 November 2025;* ***Published:*** *29 December 2025;*

Abstract

Nonperforming corporate loans pose persistent challenges for financial institutions because recoveries depend on interacting borrower, collateral, contractual, and macroeconomic conditions, while published evidence on machine learning forecasting remains fragmented. This study examines whether machine learning improves recovery rate and loss given default prediction, identifies influential determinants, and assesses validation, explainability, and institutional variation. A qualitative systematic literature review used a cross-sectional evidence snapshot and comparative case-based synthesis of published corporate loan, corporate debt, and related credit recovery investigations. Published studies, rather than cloud enterprises or newly surveyed respondents, constituted the units of analysis. Evidence was organized across six results domains and five hypotheses using structured extraction of model classes, predictor characteristics, evaluation metrics, and validation procedures, followed by qualitative appraisal informed by Statistical Learning Theory. Six illustrative publications included comparative analyses of 12 machine learning methods, 20 candidate algorithms, 768 corporate bond observations, and 2,636 defaulted corporate bonds. One United States small-business investigation examined 89,903 loan records, including 39,943 defaults; its preferred hybrid model reported mean absolute error of 0.1835, root mean squared error of 0.3108, and R^2 of 0.4577. Another investigation reported up to 30% improvement in quantile forecasting precision, while chronological validation improved out-of-time stability. Recovery predictors included collateral, borrower characteristics, loan structure, collection activity, and economic conditions. Provisional five-point evidence ratings were 3 for machine learning superiority, 4 each for predictor relevance, validation, and explainability, and 3 for contextual variation. Findings suggest conditional predictive advantages rather than universal algorithmic superiority. Because recovery definitions, financial instruments, and validation schemes differed across source studies, no pooled effects or independent significance tests were produced. For banks and credit risk managers, the synthesis supports transparent, context-sensitive recovery models, disciplined temporal validation, and careful separation of forecast accuracy from realized collection outcomes across diverse lending institutions and changing macroeconomic and legal environments.

Keywords

Machine Learning, Nonperforming Corporate Loans, Recovery Rate Prediction, Loss Given Default, Predictive Performance;

Introduction

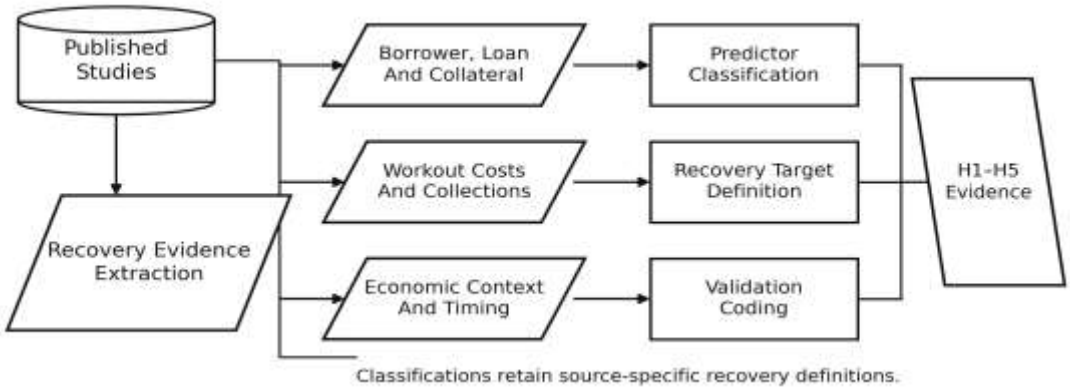
Nonperforming corporate loans refer to credit exposures extended to business enterprises for which contractual repayment obligations have become seriously overdue or repayment is considered unlikely without realizing collateral or undertaking other recovery measures. These exposures represent a significant component of banking credit risk because their financial consequences depend not only on the occurrence of default but also on the proportion of outstanding obligations ultimately recovered through restructuring, liquidation, collateral enforcement, or negotiated settlements. Recovery rate generally represents the proportion of exposure recovered after default, whereas loss given default (LGD) measures the corresponding economic loss, with the relationship between these measures commonly expressed as LGD = 1 − recovery rate when consistent recovery definitions and valuation conventions are applied (Bastos, 2010; Khieu et al., 2012). In the Bank Loan Losses-Given-Default study (pp. 1219–1243), the examination of defaulted European bank loans established the relevance of recovery timing, cumulative collections, recovery costs, and borrower characteristics in understanding realized credit losses (Dermine & Carvalho, 2006). Corporate loan recovery differs from initial default prediction because it concerns financial outcomes following deterioration in borrowers' repayment capacity and incorporates contractual seniority, collateral quality, insolvency procedures, and the duration of debt resolution. Research involving German commercial lending identified collateral coverage, borrower creditworthiness, firm size, and banking relationships as relevant characteristics of recovery outcomes (Grunert & Weber, 2009). Investigations of Italian banking exposures further demonstrated that recovery rates exhibit complex distributions that cannot always be represented adequately by conventional probability distributions (Calabrese & Zenga, 2010). The temporal characteristics of recovery processes are equally significant because the probability of successful debt resolution changes throughout the period following corporate default (Fenech et al., 2016). International evidence covering commercial loan defaults across multiple continents and more than 120 countries also demonstrates the geographical breadth and economic significance of recovery-rate analysis within banking systems (Höcht et al., 2022). Consequently, nonperforming corporate loan recovery constitutes a multidimensional financial phenomenon involving the interaction of borrower conditions, loan structures, institutional recovery arrangements, and broader economic circumstances.

The international significance of nonperforming corporate loan recovery arises from differences in financial market structures, creditor protection, insolvency regulations, economic conditions, and the institutional arrangements governing distressed corporate debt. The relationship between corporate defaults and recoveries has received substantial attention because aggregate default conditions influence the amount creditors recover from financially distressed borrowers (Altman et al., 2005). Evidence from the United States indicates that creditors experience lower recoveries when defaulted enterprises operate in industries experiencing financial distress, particularly when specialized assets have limited alternative uses and potential purchasers face liquidity constraints (Acharya et al., 2007). In Do Bankruptcy Codes Matter? (pp. 565–608), the comparative examination of France, Germany, and the United Kingdom revealed substantial differences in corporate bank-loan recoveries associated with national bankruptcy legislation, creditor rights, and collateral arrangements (Davydenko & Franks, 2008). Related evidence from Italy identified systematic relationships between macroeconomic conditions and LGD, demonstrating that the economic environment contributes to variations in realized banking losses (Caselli et al., 2008). International credit-market research has also established that default probabilities and recovery rates respond jointly to changing credit-cycle conditions, creating substantial differences in the distribution of recoveries across economic periods (Bruche & González-Aguado, 2010). Investigations of United States corporate bond markets identified relationships between recovery valuations, trading activity, liquidity conditions, contractual characteristics, and firm-level fundamentals, although market-based bond recoveries remain conceptually distinct from realized corporate bank-loan collections (Jankowitsch et al., 2014). Additional evidence from American defaulted bank loans demonstrated that the explanatory importance of recovery determinants varies across favorable and unfavorable credit conditions (Wang et al., 2020). Research conducted in Zimbabwe examined corporate bank-loan recoveries under economic downturn conditions, extending the geographical scope of recovery research to developing economies characterized by different financial and macroeconomic circumstances (Matenda et al., 2022). Collectively, these international investigations

establish substantial variation in recovery mechanisms, institutional environments, and observed outcomes across corporate lending markets.

The prediction of corporate loan recovery outcomes has traditionally relied on econometric and statistical approaches designed to estimate relationships between realized recoveries and observable financial characteristics. Common approaches include ordinary least squares regression, fractional response regression, mixture distributions, survival models, and other parametric techniques that represent the conditional behavior of recovery rates or losses given default. However, corporate recovery data frequently exhibit nonlinear associations, asymmetric distributions, heterogeneity across debt instruments, and concentrations of observations near complete recovery or substantial loss. In Comparison of Modeling Methods for Loss Given Default (pp. 2842–2855), the examination of alternative estimation techniques demonstrated that regression trees and neural networks could provide stronger predictive performance than conventional parametric approaches when model overfitting was appropriately controlled (Qi & Zhao, 2011; Sumaiya, 2026). Comparative investigations of LGD forecasting subsequently documented differences in the predictive effectiveness of statistical and computational approaches, emphasizing the relevance of model structure and evaluation procedures (Loterman et al., 2012; Mashfiqur, 2026). Additional research established that incomplete recovery histories, differences in default-resolution processes, and inconsistencies in available information can adversely affect the reliability of bank-loan LGD estimates (Gürtler & Hibbeln, 2013). Distribution-based forecasting methods addressed further complexities by representing multiple potential recovery outcomes conditional on borrower attributes, contractual characteristics, and credit conditions (Altman & Kalotay, 2014; Rony & Md, 2026). Support vector regression approaches examined in corporate bond recovery research provided evidence that nonlinear learning techniques can capture relationships overlooked by conventional specifications, particularly when differences in debt seniority are incorporated into model development (Yao et al., 2015). Japanese banking research also identified collateral, guarantees, and loan size as relevant determinants while documenting differences in predictive performance between multistage recovery models and conventional regression methods (Tanoue et al., 2017). The importance of economic uncertainty in shaping recovery distributions was further demonstrated through analyses of defaulted corporate bonds (Gambetti et al., 2019). Research on LGD model comparison additionally established that different evaluation loss functions can produce different rankings of competing forecasting models, even when identical credit exposures are examined (Hurlin et al., 2018; Jannatul, 2026). These findings illustrate the methodological diversity of recovery prediction research and the need to distinguish statistical goodness of fit from predictive reliability.

Figure 1. Recovery Prediction Evidence Pathways



Machine learning has introduced additional approaches to corporate credit-recovery forecasting through algorithms capable of identifying nonlinear associations, complex interactions, and heterogeneous relationships among borrower, contractual, institutional, and macroeconomic variables. Random forests, gradient boosting, support vector regression, artificial neural networks, regression trees, and ensemble learning represent prominent techniques investigated within the recovery-rate and LGD literature. In Ensemble Predictions of Recovery Rates (pp. 177–193), comparisons based on defaulted debt demonstrated that combining predictions from multiple models could improve recovery forecasts relative to individual

estimation methods (Bastos, 2014; Jahanara, 2026). Research concerning corporate creditor recoveries subsequently applied regularization and variable-selection techniques to identify macroeconomic information associated with recovery-rate predictions (Nazemi & Fabozzi, 2018). Additional corporate bond investigations examined multifactor support vector regression and dimensionality-reduction approaches, documenting the relevance of model specification and predictor selection to forecasting accuracy (Nazemi et al., 2018). Research directly examining nonperforming loan collections compared conventional linear approaches with beta regression and mixture-model alternatives, demonstrating the importance of accommodating complex recovery-rate distributions (Ye & Bellotti, 2019). Related investigations of defaulted corporate leasing exposures identified improvements in prediction accuracy associated with random forests and the inclusion of additional contractual information (Kaposty et al., 2020). In Forecasting Recovery Rates on Non-Performing Loans (pp. 428–444), a comparison of multiple machine learning techniques identified strong forecasting performance for Cubist models, boosted trees, and random forests while emphasizing the explanatory value of previous debt-collection activity (Bellotti et al., 2021). That investigation concerned retail credit exposures, making its results relevant to methodological comparison rather than direct evidence of corporate loan-recovery performance. More directly, comparative modeling of corporate debt LGD demonstrated that cross-validation design and the preservation of chronological relationships influence out-of-time predictive stability (Ahmed, 2026; Olson et al., 2021). Research involving defaulted corporate bonds also reported differences between traditional econometric methods and machine learning approaches under alternative temporal validation arrangements (Baumann, et al., 2022). These investigations document diverse algorithmic applications while revealing differences in target outcomes, financial instruments, predictor availability, and validation procedures.

The assessment of machine learning performance in nonperforming corporate loan recovery requires consideration of predictive accuracy, generalization, interpretability, methodological consistency, and the comparability of evidence across published banking studies. Under the generalization principles associated with Statistical Learning Theory, a model's predictive reliability concerns its ability to capture meaningful relationships rather than merely reproduce patterns within the observations used during training. Evidence from corporate LGD modeling demonstrates that validation procedures preserving temporal structure can provide more stable out-of-time forecasts than approaches that disregard chronological relationships (Olson et al., 2021; Feroz, 2025). Research examining local logit regression for loan recovery rates further demonstrates that nonlinear statistical approaches may achieve competitive predictive performance and that model rankings can differ according to the evaluation criteria applied (Sopitpongstorn et al., 2021). Interpretability represents another important dimension because the relationships identified by complex algorithms are not necessarily apparent from model structure alone. Investigations of explainable credit-loss models demonstrate how computational methods can connect nonlinear predictions to identifiable input characteristics while retaining detailed information about modeled relationships (Bastos & Matos, 2022). In Interpretable Machine Learning for Creditor Recovery Rates (article 107187), explainable machine learning was examined in the context of corporate bond recoveries, with particular attention to identifying recovery determinants and relationships that conventional approaches may not adequately reveal (Nazemi & Fabozzi, 2024). Research examining machine learning-based due diligence in secondary markets for nonperforming loans also illustrates the relevance of recovery forecasting to portfolio valuation and information asymmetry, although the underlying secured mortgage exposures differ from corporate lending portfolios (Carannante et al., 2024; Sumaiya, 2025). Existing studies therefore encompass distinct forms of credit exposure, recovery measurement, machine learning architecture, institutional context, and performance evaluation. The present research examines these differences through a systematic literature review and qualitative comparative synthesis of published findings concerning nonperforming corporate loan recovery, with particular attention to predictive performance, recovery determinants, model validation, explainability, and cross-case variation. The analysis is organized around documented empirical evidence rather than independently generated forecasting estimates, allowing reported results to be examined within their original methodological and institutional contexts.

Nonperforming corporate loans represent a significant challenge for financial institutions because delayed repayments, borrower insolvency, and uncertain recovery outcomes adversely affect credit portfolio management and financial stability. Traditional recovery forecasting approaches frequently encounter

difficulties in capturing complex relationships among borrower characteristics, collateral arrangements, economic conditions, and institutional recovery mechanisms. Machine learning has emerged as an alternative analytical approach capable of identifying nonlinear relationships and complex patterns within historical credit recovery information. However, the predictive effectiveness of different machine learning techniques varies across financial institutions, lending environments, and methodological settings. Consequently, systematically examining existing academic evidence is essential for understanding the comparative performance, reliability, and practical characteristics of machine learning approaches in forecasting nonperforming corporate loan recovery.

The prediction of nonperforming corporate loan recovery remains challenging because recovery outcomes depend on interconnected financial, contractual, institutional, and macroeconomic factors that conventional forecasting approaches may not adequately represent. Although numerous studies have investigated machine learning applications in credit risk management, the available evidence concerning corporate loan recovery prediction remains fragmented across different algorithms, financial instruments, economic environments, and validation procedures. Differences in predictive performance measures, dataset characteristics, and model explainability further complicate meaningful comparisons among published findings. Moreover, research focusing exclusively on default prediction or consumer credit recovery does not necessarily provide direct evidence concerning corporate loan recovery. The central research problem therefore concerns the absence of a sufficiently integrated understanding of machine learning predictive performance, recovery determinants, methodological reliability, and cross-context variation within the existing corporate loan recovery literature.

The primary objective of this research is to systematically examine and qualitatively synthesize existing academic evidence concerning the predictive performance of machine learning approaches in nonperforming corporate loan recovery. Specifically, the study seeks to compare machine learning algorithms with conventional statistical forecasting methods, identify borrower-specific, loan-related, collateral-related, and macroeconomic determinants associated with recovery predictions, and evaluate the influence of validation procedures, robustness, and model explainability on predictive reliability. Additionally, the research aims to investigate differences in reported forecasting performance across institutional and economic environments through comparative analysis of published corporate lending cases. Guided by Statistical Learning Theory, the study integrates these objectives into a conceptual framework that supports the systematic evaluation of existing findings and the assessment of five research hypotheses through documented scholarly evidence rather than independently generated statistical results.

RQ1: How does the predictive performance of machine learning models compare with conventional statistical approaches in forecasting nonperforming corporate loan recovery outcomes?

RQ2: Which borrower-specific, loan-related, collateral-related, and macroeconomic determinants contribute to the predictive effectiveness of machine learning models in corporate loan recovery?

RQ3: How do model validation, explainability, and differences across corporate lending environments influence the reliability and generalizability of machine learning-based recovery predictions?

This research develops five interrelated hypotheses grounded in Statistical Learning Theory and the existing literature on corporate loan recovery forecasting. The first hypothesis (H1) proposes that machine learning models demonstrate superior predictive performance compared with conventional statistical methods. The second hypothesis (H2) proposes that incorporating borrower-specific, loan-related, collateral-related, and macroeconomic determinants improves recovery prediction. The third hypothesis (H3) suggests that rigorous out-of-sample validation contributes to greater predictive reliability and generalizability. The fourth hypothesis (H4) proposes that explainable machine learning approaches enhance the transparency and interpretability of recovery predictions. Finally, the fifth hypothesis (H5) proposes that institutional characteristics, lending environments, and economic conditions influence predictive performance and generalizability. These hypotheses have been evaluated through systematic comparison of published empirical findings, considering supporting evidence, contradictory results, methodological limitations, and the overall strength of the available literature rather than conducting independent statistical hypothesis tests.

(i) Theoretical Significance: This research extends the application of Statistical Learning Theory to corporate loan recovery forecasting by examining how model complexity, predictive accuracy, validation procedures, and generalization are represented across existing machine learning studies.

(ii) Methodological Significance: The study provides a structured qualitative synthesis of published forecasting approaches, performance measures, model validation techniques, and methodological limitations, supporting more consistent interpretation of findings across heterogeneous corporate recovery studies.

(iii) Practical Significance: The research offers evidence-based insights for commercial banks, credit-risk managers, and financial institutions regarding the comparative strengths, limitations, and interpretability of machine learning approaches used in corporate loan recovery prediction.

(iv) Academic Significance: By integrating fragmented research findings and identifying inconsistencies in existing evidence, the study contributes to a clearer understanding of corporate recovery determinants, predictive reliability, and unresolved methodological questions in machine learning-based credit recovery research.

The article has been organized into an Introduction, a six-part Literature Review, a qualitative research methodology, six evidence-based Results and Findings subsections, a comparative Discussion, and final sections covering the Conclusion, Recommendations, Limitations of the Study, and APA 7th edition references.

Background and Conceptual Framework

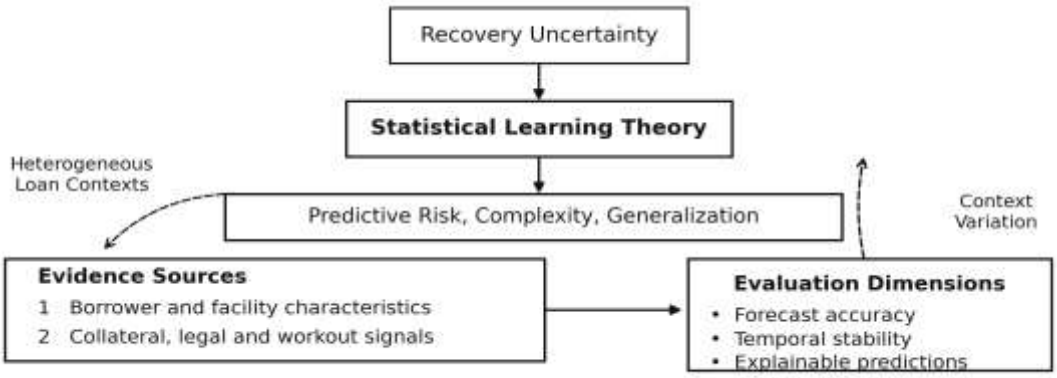
This literature review examines the evidence surrounding machine learning applications in nonperforming corporate loan recovery, emphasizing predictive performance and the conditions affecting recovery outcomes. It begins by considering the economic and institutional characteristics of distressed corporate lending and the limitations of conventional recovery forecasting. The review then evaluates machine learning approaches, relevant borrower and loan characteristics, validation practices, and model explainability. Statistical Learning Theory provides the theoretical foundation for assessing predictive accuracy, model complexity, overfitting, and generalizability across published studies. Finally, the conceptual framework organizes relationships underpinning the five research hypotheses and guides the comparative synthesis of documented evidence.

Nonperforming corporate loans emerge when business borrowers fail to meet contractual obligations or repayment becomes unlikely, redirecting lenders from scheduled cash flows toward uncertain collections. Recovery is distinct from default identification because it concerns what creditors ultimately receive, when payments arrive, and what expenditures are necessary to obtain them. Restructuring, negotiated settlements, collateral enforcement, asset sales, and insolvency proceedings can produce different financial outcomes even for borrowers with similar initial indebtedness. In Advancing Loss Given Default Prediction Models (pp. 185–230), recovery was examined across loans, bonds, and preferred shares using claim, firm, industry, country, and macroeconomic information, demonstrating the layered nature of loss estimation (Gupton, 2005; Imran & Jannatul, 2021). For corporate bank loans, this distinction requires attention to exposure at default, cash collections, recovery expenses, and the time value of money. A percentage recovered immediately does not have the same economic meaning as an equal nominal percentage collected after lengthy litigation. Default episodes themselves also vary considerably in duration and severity. In What Happens After Corporate Default? (pp. 2007–2025), evidence from Portuguese firms showed that half of corporate default episodes lasted five quarters and that renewed credit access depended partly on borrower size, default severity, and banking relationships (Ashik & Shabbir, 2022; Bonfim et al., 2012). Although that study investigated postdefault credit access rather than ultimate collection percentages, it clarifies why recovery cannot be treated as an instantaneous event. Corporate debt resolution encompasses interacting financial and institutional processes whose timing shapes both observable outcomes and the information available for predictive modeling. This distinction guides careful comparative synthesis.

Recovery dynamics also reflect the distribution of legal rights among creditors and the practical costs of enforcing those rights. Security interests, claim priority, collateral liquidity, court procedures, and debtor cooperation influence which assets become available for settlement and how much of their value remains after resolution costs. In Costs and Recovery Rates in the Dutch Liquidation-Based Bankruptcy System (pp.

105–127), average direct bankruptcy costs were reported at sixteen percent, while recovery differed markedly between firms and bank debt, illustrating how creditor position and liquidation arrangements shape realized proceeds (Couwenberg & Jong, 2008; Hasan & Adnan, 2022). These findings concern a specific national insolvency system; their numerical values should not be generalized automatically across jurisdictions. Collateral is similarly important, but its contractual valuation can differ from the amount obtained through an actual sale. Appraisals may not fully capture market liquidity, property condition, enforcement delays, or the bargaining position of distressed sellers. In The Recovery Rate for Retail and Commercial Customers in Germany (pp. 179–228), evidence from defaulted retail and small-business customers linked higher recoveries to collateral and amicable repayment agreements, while identifying systematic gaps between estimated and realized real-estate values (Ingermann et al., 2016; Akter & Sumaiya, 2022). The inclusion of retail borrowers limits direct transferability to large corporate exposures, yet the documented valuation problem remains relevant to secured lending analysis. Across these settings, recovery is best understood as a net process shaped by legal entitlements, negotiated behavior, asset realization, and the costs incurred between default and final settlement. These mechanisms complicate comparisons across banking and economic environments.

Figure 2. Learning Theory and Recovery Determinants



Corporate recovery outcomes further depend on the financial structure of the borrowing arrangement and on differences among categories of business debt. Seniority establishes the order in which creditors receive available proceeds, while covenants, guarantees, interconnected facilities, and the share of subordinated obligations affect bargaining power during distress. Leveraged loans deserve separate consideration because their borrowers often carry substantial debt burdens and their financing agreements differ from ordinary commercial lending contracts. In Modeling Recovery Rate for Leveraged Loans (pp. 231–241), a two-stage forecasting framework identified the proportion of debt ranking below a particular issuance as an important recovery determinant and found that leveraged and nonleveraged obligations exhibited different patterns of influential factors (Chen et al., 2019; Mashfiqur & Md, 2022). These differences caution against treating every corporate loan as an interchangeable observation within a single recovery distribution. Recovery amounts can also depend on whether exposure is assessed at the facility, borrower, or creditor level, since multiple claims against one enterprise may carry different priorities and collateral protections. Researchers must distinguish realized workout collections from market prices recorded shortly after default, because the latter reflect expectations and trading conditions rather than final proceeds. They must also separate recoveries on bank loans from outcomes for traded corporate bonds, which operate under different contractual and market arrangements. Accordingly, a systematic review of machine learning for nonperforming corporate loan recovery requires careful identification of recovery definitions, borrower populations, debt structures, measurement dates, and resolution contexts before comparing reported predictive performance across individual empirical studies. Such distinctions support meaningful comparison.

Conventional corporate loan recovery prediction developed around statistical estimates of loss given default and the corresponding recovery rate, using historical defaulted exposures to relate observed collections to borrower, facility, and economic characteristics. Early applications frequently relied on portfolio averages,

ordinary least squares, and generalized linear models because these methods offered transparent coefficients and relatively straightforward estimation. However, recovery proportions are bounded, commonly skewed, and often concentrated near complete loss or full repayment, creating difficulties for models that assume normally distributed errors or constant variance. In Predicting Bank Loan Recovery Rates (pp. 99–114), a mixed continuous-discrete specification combined logistic components with beta regression to accommodate extreme outcomes and variations within intermediate recovery rates using Italian banking information (Calabrese, 2014; Md & Yeasin, 2022). This approach demonstrated how conventional statistical modeling can address distinctive recovery distributions without depending on complex machine learning algorithms. In Estimating Loss Given Default Based on Beta Regression (pp. 3329–3344), a Jordanian banking portfolio was examined through beta regression with separate mean and precision components, incorporating institutional financial indicators and macroeconomic variables (Jaber et al., 2021; Ashfaq & Mohammad, 2022). Its comparison of alternative link functions emphasized that statistical specification affects estimated relationships and model fit. These approaches provide interpretable baselines for comparing algorithms, although their reported goodness of fit does not automatically establish reliability for new corporate defaults. Moreover, samples differing in collateralization, borrower composition, recovery definitions, and observation periods can yield contrasting results under an identical regression family. Conventional recovery forecasting therefore requires attention to distributional assumptions and consistent outcome measurement across financial institutions.

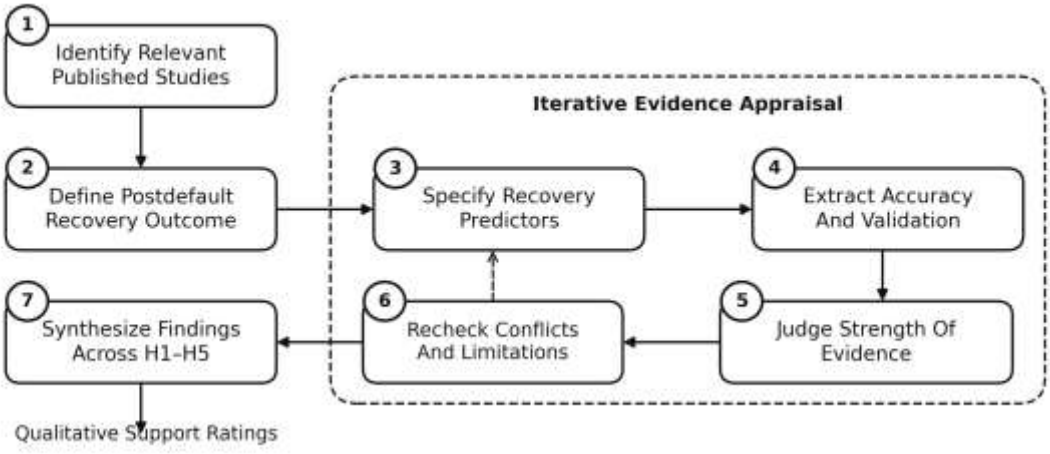
Statistical approaches also differ according to the information recorded during collection, particularly when a troubled corporate exposure remains unresolved for an extended period. A model estimated entirely from borrower accounts at default can overlook developments arising from collateral enforcement, negotiation, and legal proceedings, whereas a periodically updated specification can incorporate the changing recovery process. In Effects of Debt Collection Practices on Loss Given Default (pp. 21–31), a model based on indicators of legal and internal collection actions performed better than a conventional firm-characteristics model (Han & Jang, 2013; Imran & Jannatul, 2022). The finding identifies collection activity as a relevant source of explanatory information, not merely a consequence observed after recovery has occurred. Another conventional approach is Tobit regression, which accommodates censoring but requires careful justification when recovery values cluster at boundaries for economic rather than observational reasons. In Loan Level Loss Given Default Study (pp. 168–177), Indian bank records covering two decades supported a multifactor Tobit analysis of loss outcomes across sectors, facility types, and collateral categories (Bandyopadhyay, 2022; Morshedul & Islam, 2022). That evidence also connected recovery performance with security quality, liquidation considerations, and negotiated settlements. Together, these studies indicate that interpretability depends as much on variable definitions and the point at which information becomes available as on the choice of regression equation. For a systematic review, reported performance must therefore be distinguished between models based solely on origination information, those updated after default, and those calculated after final resolution. Otherwise, an apparent forecasting advantage may reflect richer information rather than a more effective predictive method.

Conventional forecasting is not restricted to a single conditional average, because lenders also face uncertainty concerning extreme losses and the overall distribution of possible recoveries. Parametric mixture models, censored regressions, and Bayesian specifications attempt to represent features that simpler linear estimates cannot capture adequately, especially observations concentrated at zero or one. In Using the Bayesian Sampling Method to Estimate Corporate Loss Given Default Distribution (article 101540), posterior simulation with a zero-one-inflated beta specification reproduced corporate loss distributions more closely than a simpler Bayesian Tobit alternative in the examined evidence (Kaniz & Jahanara, 2022; Zhang & Zhao, 2024). This comparison demonstrates that substantial modeling flexibility is available within statistical frameworks traditionally distinguished from machine learning. Accordingly, categorical claims that computational algorithms outperform conventional prediction are difficult to assess when the statistical benchmark is restricted to an elementary linear equation. Model assessment also depends on whether a study predicts the conditional mean, captures an entire loss distribution, or evaluates tail outcomes relevant to stress testing. A method selected for transparent coefficient interpretation may be judged differently from one developed for accurate portfolio-wide risk quantification. Recovery timing creates additional problems because completed cases can differ systematically from unresolved cases, potentially distorting the available

historical sample. Comparisons across studies consequently require clear identification of observed outcomes, sample composition, calibration periods, validation design, and error measures. Within the present literature-based inquiry, conventional statistical models constitute a diverse reference category against which published machine learning performance can be interpreted without introducing newly estimated analytical values or assuming identical lending conditions.

Machine learning approaches to nonperforming corporate loan recovery prediction seek to represent relationships that fixed statistical specifications may overlook, particularly when borrower finances, collateral conditions, recovery actions, and economic conditions interact. Their purpose is to estimate recoverable proportions or losses after default rather than the probability that a performing borrower will default. Support vector regression, regression trees, and hybrid procedures differ in how they accommodate nonlinearities, partition observations, and combine financial indicators. In Forecasting Loss Given Default Models (pp. 376–392), comparisons involving corporate and home equity loans examined nonlinear support vector regression, regression trees, transformed regression, and hybrid models against ordinary least squares. Predictive rankings differed between out-of-time and out-of-sample evaluations, while macroeconomic indicators improved forecasting performance in the examined settings (Stabak & Shahid, 2022; Tobback et al., 2014). This finding links algorithm selection to validation design and the changing conditions under which a corporate exposure is resolved. Predictor construction is equally important because repayment behavior and credit relationships may convey information unavailable in balance sheet ratios alone. In Bank Loans Recovery Rate in Commercial Banks (pp. 139–172), monthly observations concerning nonfinancial corporations were analyzed through Bayesian model averaging and quantile regression, incorporating financial and behavioral records and distinguishing point-in-time from through-the-cycle perspectives (Nehrebecka, 2019; Sadia & Mostari, 2022). Although these are not interchangeable with tree-based algorithms, their examination clarifies which recovery predictors require evaluation across modeling families. Together, these studies position credit history, economic indicators, and observation timing as essential considerations for comparing reported machine learning recovery forecasts. They also identify limits to uncomplicated cross-study performance comparisons.

Figure 3. Iterative Recovery Evidence Coding Process



Ensemble learning extends recovery forecasting by combining distinct prediction rules rather than relying on one functional relationship. Tree-based models partition cases according to observed characteristics, support vector regression estimates flexible nonlinear relationships, and fusion procedures combine outputs from alternative learners. Such flexibility is potentially useful when seniority, collateral, industry conditions, and macroeconomic shocks jointly shape the amounts creditors obtain after corporate distress. In Fuzzy Decision Fusion Approach for Loss-Given-Default Modeling (pp. 780–791), researchers combined regression, support vector methods, and decision trees through an optimized fuzzy framework for corporate bond losses. They also incorporated principal components from extensive macroeconomic information and reported improved predictive accuracy relative to individual benchmark methods (Nazemi et al., 2017; Feroz & Pervaz, 2022). Corporate bonds provide closely related evidence about creditor recovery, although market-based debt measures should not be equated with realized collections on nonperforming bank loans.

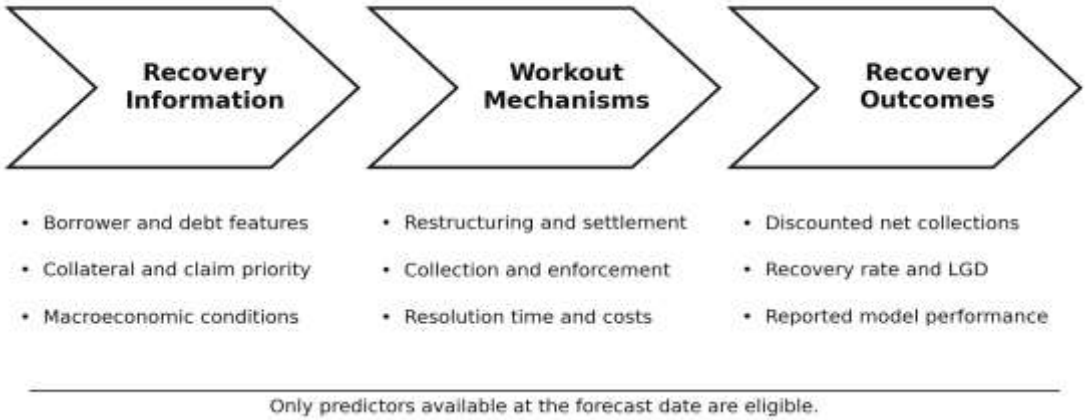
Model combination does not necessarily mean indiscriminate inclusion of every available variable. In Meta-Learning Approaches for Recovery Rate Prediction (article 124), a comparison involving twenty candidate algorithms showed the value of combining models with security-specific characteristics and a limited set of theoretically relevant economic and uncertainty indicators (Mohammad, 2023; Gambetti et al., 2022). This distinction places feature relevance alongside algorithmic complexity when interpreting reported performance. For corporate banking applications, plausible information categories include claim priority, collateral support, borrower characteristics, recovery procedures, and prevailing economic conditions. A qualitative review must therefore compare predictor availability and recovery definitions before attributing performance differences to an algorithm itself, because related corporate debt instruments can encode substantially different recovery processes.

Neural architectures add another modeling perspective by learning layered transformations of recovery predictors and allowing conditional relationships to vary across portions of the loss distribution. This distinction matters because predicting the average loss may conceal the risk of exceptionally poor recoveries on distressed corporate exposures. In Opening the Black Box: Quantile Neural Networks for Loss Given Default Prediction (article 106334), a neural extension of linear quantile regression captured nonlinear effects and interactions in bank loan losses. The authors reported improvements of up to thirty percent in quantile forecast precision relative to their benchmark, with particular gains in higher loss quantiles (Adnan & Iftekhar, 2023; Kellner et al., 2022). Such results concern the study's own setting and evaluation measures rather than a universal improvement applicable to every corporate loan portfolio. The methodological significance extends beyond whether a particular algorithm has the lowest reported error. Quantile forecasts, mean recovery estimates, and classifications of complete versus partial collection answer different questions and cannot be ranked through one undifferentiated accuracy claim. Furthermore, deep models depend on the quality and timing of input records, the number of resolved defaults, and safeguards separating training information from subsequent recovery outcomes. For a qualitative systematic review, explanatory indicators should therefore be organized by borrower, loan, collateral, collection, and economic categories, while predictive results should be compared only when the target measure and validation approach are sufficiently compatible. This preserves the distinction between demonstrated algorithmic flexibility and verifiable performance in actual corporate loan recovery settings. These findings cannot establish neural superiority across institutions.

Mechanisms and outcome domains

The conceptual framework for nonperforming corporate loan recovery links the selection of forecasting methods with the availability and relevance of borrower, facility, collateral, collection, and macroeconomic information. Its central outcome is the recovery rate after default, interpreted consistently with the definition adopted by each primary study. For a workout-based corporate loan, the relationship can be represented conceptually as the discounted value of net cash collections relative to exposure at default, and loss given default as the complement of recovery rate. Here exposure at default is EAD, collections are C, recovery costs are K, the discount rate is r, and the resolution horizon is T. This equation defines the target; it is not applied to calculate new recovery values. At the prediction level, the model describes how borrower, loan, collateral, workout, and macroeconomic predictors may contribute to estimated recovery. In Weighted Quantile Regression Forests for Bimodal Distribution Modeling (article 545), an ensemble approach addressed complex loss distributions and showed stronger predictive stability than selected alternatives in a Polish banking dataset (Ahmed & Sadia, 2023; Gostkowski & Gajowniczek, 2020). The first hypothesis (H1) therefore concerns comparative predictive performance rather than automatic algorithmic superiority. The second hypothesis (H2) concerns the added informational contribution of relevant recovery determinants. In Exploring Industry-Distress Effects on Loan Recovery (article 6), industry distress, collateral status, and debt structure were examined across defaulted debt and loan instruments using quantile-based analysis (Islam & Md, 2023; Chuang & Chen, 2023). These findings connect the framework's predictor categories to heterogeneous recoveries, while differences between traded debt and corporate bank loans remain explicit in evidence classification. Crucially, predictor availability depends on the forecast date, since collection information observed after final resolution cannot justify an earlier prediction.

Figure 4. Recovery Mechanisms and Outcome Domains



$$RR_i = \frac{1}{EAD_i}\sum_{t=1}^{T_i}\frac{C_{it}-K_{it}}{(1+r_i)^t}, \quad LGD_i = 1 - RR_i$$

$$\widehat{RR}_{i,m} = f_m(B_i, L_i, S_i, W_i, M_i)$$

The framework places validation quality and explanatory transparency between model development and credible interpretation of forecasting results. Statistical Learning Theory distinguishes the ability to fit recorded defaults from the ability to forecast recoveries for observations occurring outside the estimation sample. Accordingly, the generalization gap introduced in Section 2.5 remains the principal theoretical criterion for appraising reported validation practices. A smaller gap is informative only when training and external losses use compatible definitions and genuinely separated information. In Modeling Ultimate Loss-Given-Default and Time-to-Resolution on Corporate Debt (pp. 426–459), a joint consideration of recovery severity and resolution duration addressed the bias introduced when defaults had not reached final settlement (Asheke, 2023; Jacobs, 2024). This evidence supports examining censoring, observation horizons, and chronological testing when assessing the third hypothesis (H3) concerning reliability. Explainability constitutes a separate dimension of model assessment because feature importance cannot establish that a forecast generalizes or that a relationship is causal. In Applications of Explainable Artificial Intelligence in Finance (pp. 867–907), a systematic review distinguished intrinsically interpretable methods from post hoc explanations and classified their functions across financial applications (Yeasin, 2023; Weber et al., 2024). That distinction frames the fourth hypothesis (H4), which evaluates whether published recovery studies provide understandable explanations of particular predictions and influential variables. The conceptual framework consequently treats predictive accuracy, external validity, and transparency as complementary evidence dimensions, not interchangeable measures of success. All judgments remain confined to documented findings in the included publications. The review also checks whether studies clearly distinguish information available at default from later workout developments.

Institutional and economic circumstances form the contextual layer of the conceptual framework, potentially altering both actual corporate recovery outcomes and the stability of forecasting relationships. Insolvency procedures, collateral enforcement, contract priority, sectoral distress, and changes during lengthy workouts can shape the distributions from which learning algorithms obtain their observations. In Systematic Effects Among Loss Given Defaults and Their Implications on Downturn Estimation (pp. 1113–1144), comparisons of banking recovery information from the United States and Europe identified regional differences in systematic loss effects that did not simply mirror conventional economic cycles (Betz et al., 2018; Jahanara & Kaniz, 2023). This evidence grounds the fifth hypothesis (H5), under which institutional and economic variation is considered a condition affecting comparability and generalizability rather than an independently estimated causal effect. Taken together, the conceptual relationships align the five hypotheses with five evidence dimensions: whether machine learning exceeds conventional benchmarks, whether informative predictors improve forecasts, whether sound validation supports stability, whether explanations clarify model behavior, and whether context changes reported results. Evidence extraction will record the recovery definition, instrument type, borrower population, predictor timing, competing

algorithms, test design, interpretation procedures, and institutional setting for each eligible study. The synthesis will then compare reported findings within compatible categories and classify hypothesis support as favorable, mixed, contradictory, or insufficient. No coefficients, pooled effect sizes, or model errors will be generated from unavailable primary data. The framework therefore organizes a qualitative systematic comparison, while the recovery and generalization equations remain conceptual criteria used to interpret published evidence consistently. This prevents unsupported confirmation claims.

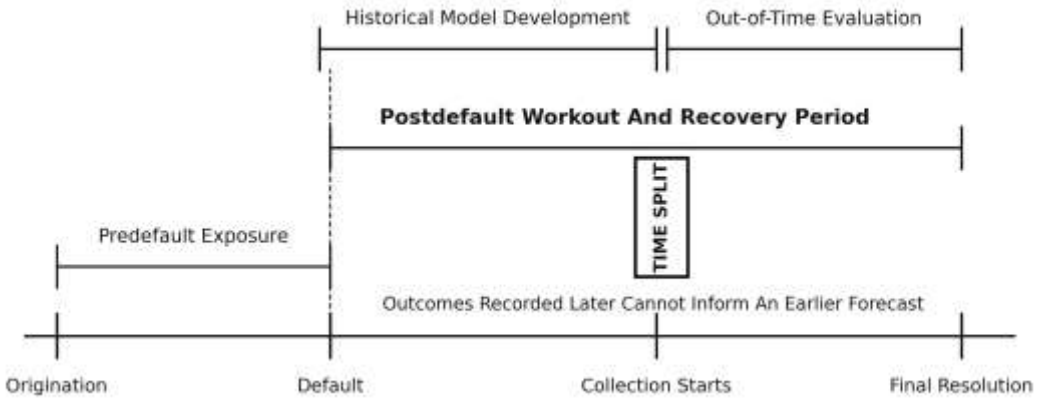
Measurement concepts

Predictive performance in nonperforming corporate loan recovery research concerns how accurately an estimated model reproduces recovery outcomes for exposures that were not used in its development. Consequently, a low error within the estimation sample cannot establish that a forecasting technique will perform reliably on new defaults. In Loss Given Default for Leasing (pp. 364–375), comparisons of regression trees, model trees, mixture specifications, and conventional benchmarks showed that approaches fitting historical observations particularly well could perform substantially worse outside their training sample (Hartmann-Wendels et al., 2014; Sheak & Tahmedur, 2023). Although leasing differs contractually from corporate bank lending, this evidence makes the distinction between fitting and forecasting particularly clear. Predictive accuracy must also be interpreted relative to the observed shape of the loss distribution. Recovery outcomes frequently contain concentrations near the boundaries, asymmetry, and heterogeneous patterns across banking portfolios. In Heterogeneities Among Credit Risk Parameter Distributions (pp. 251–287), an investigation of defaulted loans across sixteen European countries compared twenty estimation methods and found that preferred approaches changed with the modality of the underlying loss distributions (Gürtler & Zöllner, 2023; Kaium, 2023). The result challenges the assumption that one complex algorithm should dominate every recovery portfolio. For qualitative cross-study comparison, reported mean absolute error, root mean squared error, and comparable forecasting criteria should be examined alongside portfolio composition, outcome definitions, and test design. Differences in such conditions can generate contrasting model rankings without establishing a universal superiority of one method. A credible comparison therefore evaluates how performance was measured before considering which algorithm appeared strongest across distinct portfolios.

Model validation constitutes a separate dimension of predictive reliability because recovery models are estimated from defaults occurring under particular economic and institutional conditions. Randomly dividing observations into training and testing sets may demonstrate discrimination within one dataset while providing limited evidence that performance persists across subsequent credit cycles. Such a limitation matters when corporate restructurings, collateral valuations, and economic downturns change the relationship between loan characteristics and ultimate losses. In Downturn LGD Modeling Using Quantile Regression (pp. 42–56), research based on defaulted small and medium enterprise loans emphasized assessing the entire loss distribution and demonstrated why conventional comparisons of conditional means can miss important information about extreme losses (Krüger & Rösch, 2017; Pervaz, 2023). Validation should therefore match the prediction objective, distinguishing average recovery estimates from forecasts of unusually severe losses. The choice of error measure can also change the relative assessment of competing models when larger exposures receive greater weight. In Deep Learning for Modeling the Collection Rate (pp. 240–252), an investigation of purchased defaulted consumer receivables reported that deep learning performed best under exposure-weighted evaluation, whereas unweighted comparisons showed similar performance for boosting and deep learning (Badrul & Mominul, 2023; Rezazadeh, et al., 2022). This evidence is methodological rather than direct confirmation for corporate loans because the underlying receivables originated in telecommunications. Taken together, the reviewed findings show why validation periods, sample separation, loss distributions, and exposure weighting must be identified before published estimates are treated as comparable. Model rankings without this information remain difficult to interpret across corporate lending contexts or historical periods. Explainability addresses an additional requirement of recovery prediction: financial decision makers need to understand the information associated with estimated losses, not merely obtain a numerical forecast. Complex algorithms can accommodate interactions among contractual seniority, collateral, company characteristics, and macroeconomic conditions, yet their relationships may remain obscure when only overall accuracy is reported. Feature attribution, sensitivity analysis, and locally interpretable predictions can help distinguish

whether a model is responding to economically plausible information or exploiting incidental patterns in historical cases. Nevertheless, explanatory displays should not be treated as evidence that identified predictors causally determine eventual collections.

Figure 5. Recovery Timing and Validation Windows



In Quantifying Uncertainty of Machine Learning Methods for Loss Given Default (article 1076083), an application of deep evidential regression to defaulted corporate bonds separated uncertainty associated with the data from uncertainty attributable to model knowledge and applied explainable artificial intelligence to examine important drivers (Nagl et al., 2022; Nishat, 2023). The authors found that uncertainty arising from the data was markedly greater, highlighting limits to the precision of individual recovery estimates. Since their observations concerned market-based bond losses, the findings require careful distinction from discounted workout recoveries on nonperforming corporate loans. For the present review, interpretability, uncertainty quantification, and robustness are therefore assessed as related but distinct attributes. Transparent feature explanations do not substitute for out-of-time validation, while accurate forecasts alone do not establish that a model is comprehensible or stable. Evidence synthesis must examine the original outcome, explanation method, validation approach, and documented limitations together when assessing the trustworthiness of reported predictive performance across banking environments.

Statistical Learning Theory provides the principal theoretical framework for evaluating whether machine learning algorithms predict nonperforming corporate loan recoveries beyond the observations used to construct them. Rather than equating statistical fit with forecasting success, the theory treats learning as the selection of a prediction function capable of limiting expected error for previously unseen cases. Recovery outcomes are particularly demanding because the relationships connecting borrower finances, collateral, contractual priority, and economic conditions can be nonlinear and unstable across periods. In Another Look at Statistical Learning Theory and Regularization (pp. 958–969), the distinction between function approximation and predictive risk minimization explains why increasing model flexibility does not automatically improve generalization and why complexity control remains essential when financial observations are limited (Cherkassky & Ma, 2009; Sadia, 2023). For this review, the central theoretical expression is the generalization gap, where the out-of-sample expected risk is compared with the empirical training loss. Here, the loss function represents absolute recovery prediction error, while the gap expresses the difference between training and genuinely external performance. In Loss Given Default Estimation (pp. 19–37), comparisons involving Japanese regional banks used cross-validation and out-of-time testing to examine whether nonlinear, boosted, and support vector approaches retained predictive advantages beyond development samples (Stabak & Siam, 2023; Tanoue & Yamashita, 2019). The equations therefore serve as interpretive standards for published validation evidence, not as statistics calculated from newly collected loan data. They distinguish variation in recovery mechanisms from uncertainty produced by sample-specific fitting, particularly when distressed borrowers enter workouts during different economic periods and under different institutional rules. Such distinctions matter across studies.

$$\mathcal{G}(f) = \mathcal{R}_{\text{out}}(f) - \hat{\mathcal{R}}_{\text{train}}(f)$$

$$\mathcal{R}_{\text{out}}(f) = \mathbb{E}\big[\ell\big(Y, f(X)\big)\big]$$

$$\hat{\mathcal{R}}_{\text{train}}(f) = \frac{1}{n}\sum_{i=1}^{n} \ell\left(y_i, f(x_i)\right), \quad \ell(y, \hat{y}) = |y - \hat{y}|$$

The theory further emphasizes a tradeoff between representational complexity and the quantity and quality of available information. Flexible algorithms can reproduce intricate recovery patterns, yet they may also capture accidental correlations arising from small samples, unusual insolvency episodes, or imperfectly recorded collateral outcomes. In this setting, an algorithm with excellent training accuracy may provide a weaker estimate of eventual collections than a simpler benchmark evaluated under comparable conditions. In Comparison Study of Two-Step LGD Estimation (pp. 155–177), alternative probability machines were used in a two-stage loss given default framework, and random forests performed best for its initial classification component relative to the examined alternatives (Feroz, 2023; Tanoue et al., 2020). That result should be interpreted for the specific prediction stage and portfolio analyzed, rather than generalized to every corporate recovery task. Model structure also matters because recovery can involve curing, partial collection, or write-off, which do not necessarily follow one homogeneous statistical process. In Loss Given Default Decomposition (pp. 1187–1199), modeling these in-default events jointly produced favorable out-of-sample comparisons against two benchmarks while retaining meaningful information about collection activities (Starosta, 2021; Tahmedur & Rifat, 2023). From a learning-theoretic perspective, this decomposition changes the prediction task itself and may reduce uncertainty by representing distinct recovery pathways explicitly. The present review consequently examines model objectives, predictor timing, sample size, training procedures, and external tests together. Without these distinctions, favorable published errors cannot be attributed securely to algorithmic sophistication, and numerical rankings drawn from unrelated financial portfolios remain methodologically fragile. Comparisons also require equivalent targets and documented measurement periods.

Applying Statistical Learning Theory within a qualitative systematic review requires evaluating reported model generalization without reproducing algorithms, reconstructing datasets, or pooling incompatible error measures. The proposed generalization-gap expression provides the consistent interpretive lens: stronger evidence occurs when a study reports independent validation and documents small deterioration relative to its own training benchmark, while weaker evidence arises from training fit alone. This assessment must remain conditional because historical recovery cases may be affected by temporal shifts, incomplete collections, censoring, and uneven distributions of exposure amounts. In Assessing the Loss Given Default of Bank Loans Using the Hybrid Algorithms Multi-Stage Model (article 505), experiments on United States Small Business Administration credit records compared multistage hybrid methods with competing routes and examined the stability of findings after introducing business and macroeconomic features (Mohammad, 2024; Fan et al., 2023). The underlying small-business portfolio offers relevant recovery-modeling evidence, although its characteristics should not be assumed identical to large corporate distressed-loan portfolios. For the five review hypotheses, generalization is pertinent to comparisons with traditional methods, the contribution of predictors, rigorous validation, explainability, and institutional variation. Evidence will therefore be classified according to whether performance claims depend on training metrics, randomized holdouts, temporally separated observations, or tests across portfolios. Model explanations will also be considered alongside predictive results because interpretable variables do not necessarily guarantee robust forecasts. This framework makes the theoretical equation usable throughout the study as a qualitative appraisal criterion, without manufacturing new analytical values or claiming that unreported prediction gaps have been independently and directly measured.

Methods

Design and reporting

The study has adopted a qualitative systematic literature review with a cross-sectional evidence snapshot and comparative published-case orientation to examine recovery forecasting rather than originate new loan-level estimates.

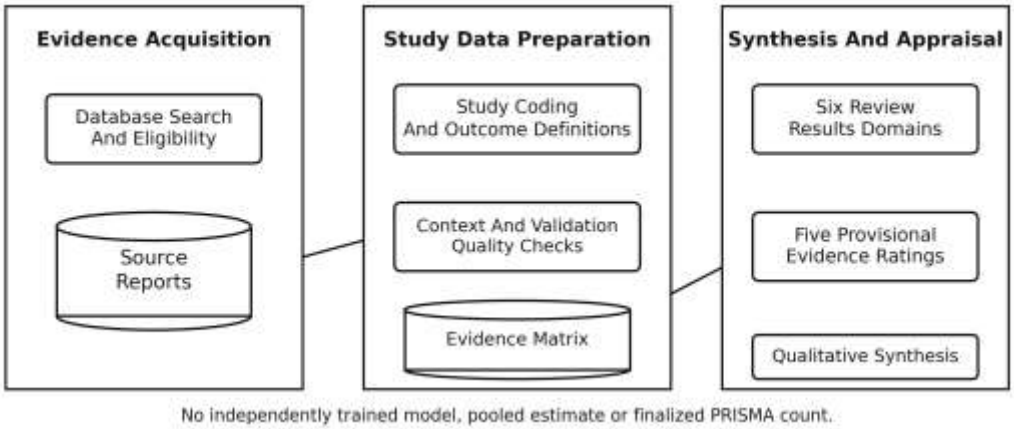
Eligibility criteria

Its case-study context has comprised documented corporate bank-loan, corporate debt, and explicitly distinguished adjacent credit-recovery settings, with published studies and their reported model comparisons serving as units of analysis instead of bank staff or survey respondents.

Information sources and search

Literature search procedures have considered Scopus, Web of Science, ScienceDirect, and targeted searches of peer-reviewed financial research using combinations of machine learning, nonperforming loans, loss given default, recovery rate, corporate lending, prediction, validation, and explainability; a finalized reproducible search log and PRISMA counts have not been supplied.

Figure 6. Qualitative Review and Synthesis Workflow



Selection process

Purposeful eligibility criteria have prioritized research on postdefault recovery rates and loss given default, and have distinguished bank-loan recoveries from bond market pricing, consumer debt collection, or prediction of default occurrence alone.

Data extraction

The extraction instrument has been defined as a structured study-coding form recording authorship, year, population, recovery definition, model class, predictors, original reported performance metrics, validation arrangements, and case context. Microsoft Word has supported manuscript preparation and table formatting, Microsoft Excel may support study logs and reference extraction, and EndNote may be used for bibliography management; SPSS has not been used to estimate regressions or survey statistics because no original dataset has been assembled.

Appraisal of the included studies

A five-point evidence rubric has been described for qualitative appraisal (1 = strongly contradictory, 2 = contradictory, 3 = mixed, 4 = supportive, 5 = strongly supportive), not as a questionnaire administered to human participants. No respondent recruitment, population survey, empirical pilot test, reliability alpha, or psychometric validity test has been undertaken; instead, source relevance and methodological transparency have been considered in the planned critical appraisal, and any initial coding rubric remains provisional until independently checked.

Synthesis

Evidence has been synthesized thematically across six results domains and provisionally mapped to five hypotheses, with Statistical Learning Theory guiding assessment of out-of-time performance and generalization. The synthesis has used numerical values only when reported within the cited individual publications and has not pooled incompatible results into a formal meta-analysis.

Results

The results synthesize published research examining machine learning approaches to nonperforming corporate loan recovery, emphasizing predictive accuracy, recovery determinants, validation reliability, interpretability, and institutional differences. The evidence has been organized around five hypotheses and three research questions, with direct corporate loan studies distinguished from related investigations of corporate bonds, leasing exposures, and small-business lending. In a corporate debt study, 12 machine learning methods were evaluated, and preserving chronological order during cross-validation produced more stable out-of-time forecasts, emphasizing the importance of realistic testing procedures (Ahmed, 2024; Olson et al., 2021). Another corporate bond recovery study compared 20 candidate algorithms and identified advantages in combining forecasting methods with security-specific information and economic indicators (Gambetti et al., 2022; Pervaz, 2025). Research using 89,903 United States small-business loan records, including 39,943 defaults, reported a mean absolute error of 0.1835, a root mean squared error of 0.3108, and an R^2 of 0.4577 for its best evaluated hybrid configuration (Islam, 2024; Fan et al., 2023). These figures describe that study alone and cannot be interpreted as pooled corporate loan results. Another bank-loan investigation reported improvements of up to 30% in quantile forecasting precision and found that macroeconomic predictors were up to twice as influential in American observations as in European observations (Asheke, 2024; Kellner et al., 2022). Collectively, these findings offer qualified evidence relevant to H1 and H2, while the reported validation comparisons inform H3. Published explainability analyses illuminate H4 by examining which characteristics influence model outputs, although interpretability does not itself demonstrate greater forecasting accuracy. Differences between national markets, debt instruments, economic conditions, and prediction horizons inform H5 and restrict direct cross-study comparisons. The evidence addresses the study objectives through comparative synthesis rather than original numerical estimation. Hypothesis judgments remain provisional until all eligible studies have been systematically selected, appraised, and synthesized, and no pooled effect sizes or independent significance tests are claimed.

The qualitative five-point evidence rubric has been interpreted as 1 = strongly contradictory, 2 = contradictory, 3 = mixed, 4 = supportive, and 5 = strongly supportive. Such ratings have not represented a respondent-based Likert survey or psychometric measurement; they have been judgments about the direction and strength of published evidence, which have remained provisional in the absence of finalized PRISMA counts and comprehensive quality appraisal.

Characteristics of the included studies

Table 1: Characteristics of Published Machine Learning Recovery Studies

Study	Research context	Principal variable	Published numerical evidence	Principal finding
Olson et al. (2021)	U.S. corporate debt	ML and validation	12 machine learning methods	Chronological validation improved out-of-time stability.
Bellotti et al. (2021)	European nonperforming loan collections	Recovery forecasting algorithms	Numerical performance comparison reported	Tree-based and rule-based models performed favorably.
Kellner et al. (2022)	U.S. and European bank loans	Quantile forecasting and explainability	Up to 30% improvement	Quantile neural networks improved forecast precision.
Gambetti et al. (2022)	Corporate bond recovery	Meta-learning and predictors	20 algorithms; 768 observations	Selected-predictor combinations performed favorably.
Fan et al. (2023)	U.S. small-business loans	Hybrid LGD prediction	89,903 records; 39,943 defaults	Hybrid multistage models provided competitive predictions.
Nazemi and Fabozzi (2024)	U.S. corporate bonds	Interpretable machine learning	2,636 defaulted bonds	Interpretable models showed favorable out-of-time results.

Table 1 has presented six publications that have provided evidence directly or indirectly relevant to machine learning-based recovery prediction. The studies have represented different financial instruments, lending environments, prediction objectives, and validation procedures. Olson et al. (2021) have provided corporate debt evidence through comparisons involving 12 machine learning methods, while Gambetti et al. (2022) have examined 20 algorithms using 768 corporate bond observations. Fan et al. (2023) have analyzed 89,903 small-business credit records, including 39,943 defaults. Nazemi and Fabozzi (2024) have investigated 2,636 defaulted corporate bonds. These sample sizes have described separate original investigations and have not been combined into a single research dataset. From the perspective of Statistical Learning Theory, the studies have demonstrated why predictive generalization must be evaluated within the particular population and observation period from which each model has learned. The evidence has established a preliminary basis for addressing the research objectives, but the six publications have not been represented as a complete PRISMA-selected sample.

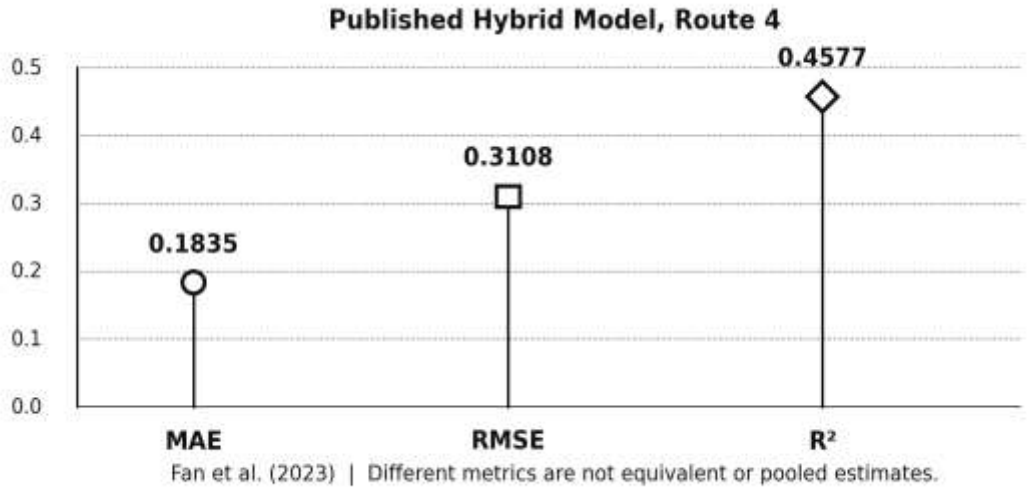
Machine learning models and predictive performance

Table 2: Comparative Predictive Performance of Machine Learning Approaches

Study	Prediction method	Performance variable	Reported result	Hypothesis
Olson et al. (2021)	Multiple ML models	Out-of-time stability	12 methods; chronological validation improved stability	H1, H3
Bellotti et al. (2021)	Cubist, boosted trees, random forests	Recovery prediction accuracy	Rule-based/tree-based models exceeded several alternatives	H1
Gambetti et al. (2022)	Meta-learning ensembles	Comparative prediction	20 candidate algorithms evaluated	H1, H2
Kellner et al. (2022)	Quantile neural networks	Quantile forecast precision	Up to 30% improvement against benchmark	H1, H4
Fan et al. (2023)	Hybrid multistage Route 4	MAE	0.1835	H1
Fan et al. (2023)	Hybrid multistage Route 4	RMSE	0.3108	H1
Fan et al. (2023)	Hybrid multistage Route 4	R^2	0.4577	H1

Table 2 has summarized the principal predictive performance findings associated with the first research objective and H1. The reviewed literature has identified favorable forecasting outcomes for several machine learning approaches, including ensemble learning, tree-based prediction, and neural networks. Fan et al. (2023) have reported an MAE of 0.1835, RMSE of 0.3108, and R^2 of 0.4577 for the preferred Route 4 configuration of their hybrid multistage model. Kellner et al. (2022) have reported improvements reaching 30% in quantile forecasting precision relative to their benchmark. These results have demonstrated performance advantages within their respective datasets, although they have not established a universal ranking across corporate lending environments. Statistical Learning Theory has provided an interpretive foundation by distinguishing predictive accuracy achieved during model development from performance obtained on independent observations. Consequently, H1 has received conditional support from the comparative evidence. The findings have suggested that machine learning can outperform conventional alternatives under particular conditions, while differences in portfolio composition, validation methods, and recovery definitions have prevented definitive generalization.

Figure 7. Published Hybrid Model Performance Indicators



Corporate loan recovery determinants

Table 3: Recovery Predictors and Published Findings

Predictor category	Study	Variable examined	Published result	Hypothesis
Collateral	Fan et al. (2023)	Real estate mortgage indicator	Absolute OLS coefficient: 1.629	H2
Loan characteristics	Fan et al. (2023)	Portion variable	Absolute OLS coefficient: 0.590	H2
Macroeconomic	Fan et al. (2023)	Recession indicator	Absolute OLS coefficient: 0.1778	H2
Macroeconomic importance	Kellner et al. (2022)	U.S. versus European macro effects	Up to twice as important in U.S. observations	H2, H5
Security-specific	Gambetti et al. (2022)	Seniority, security and economic indicators	Selected predictors strengthened forecasting	H2
Collection history	Bellotti et al. (2021)	Previous debtor contact and repayment	Recovery variables enhanced forecasting	H2

Table 3 has examined the determinants associated with recovery forecasting and has addressed the second research objective. The reviewed evidence has identified collateral arrangements, loan-specific characteristics, macroeconomic conditions, and historical recovery activity as relevant sources of predictive information. Fan et al. (2023) have reported absolute OLS coefficients of 1.629 for their real estate indicator, 0.590 for the Portion variable, and 0.1778 for the recession indicator in the third stage of their model. These coefficients have represented findings from that particular specification and have not been interpreted as standardized effects or causal estimates. Kellner et al. (2022) have documented differences in macroeconomic feature importance between American and European lending observations. Bellotti et al. (2021) have additionally demonstrated that information about previous collection activity can contribute to recovery prediction. Statistical Learning Theory has explained the relevance of these findings through the relationship between informative predictors and forecasting error. Accordingly, the literature has provided supportive evidence for H2, while differences in measurement and predictor availability have limited direct comparisons across studies.

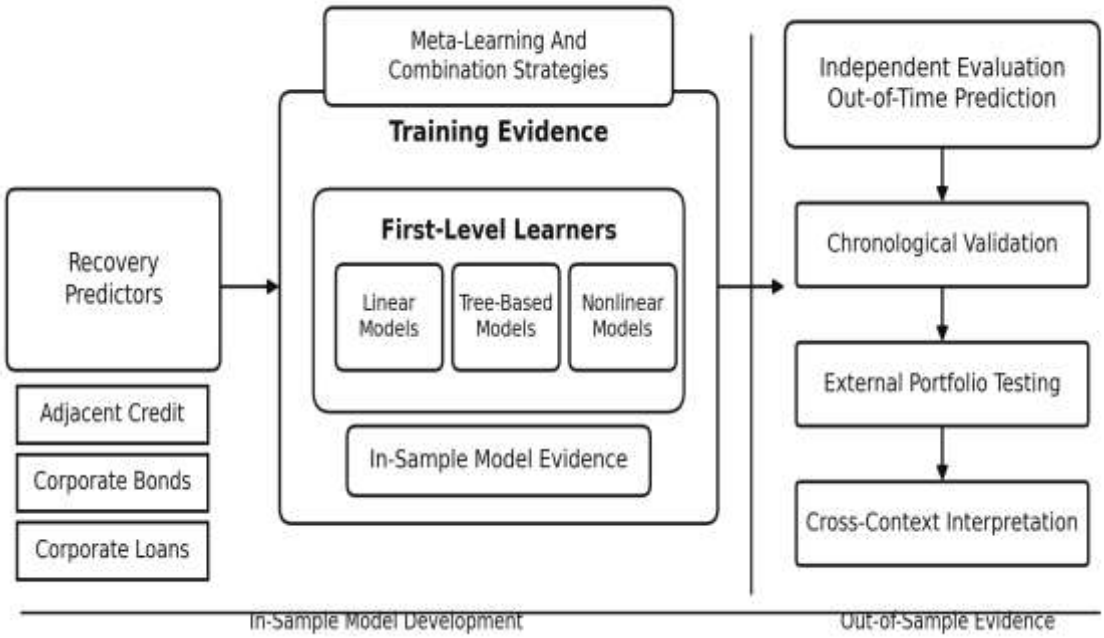
Model validation, robustness, and explainability

Table 4: Validation Procedures and Interpretability Findings

Study	Assessment dimension	Reported numerical evidence	Principal finding	Hypothesis
Olson et al. (2021)	Cross-validation	3 validation schemes	Chronological preservation improved out-of-time stability	H3
Kellner et al. (2022)	Robustness and feature interpretation	Up to 30% precision improvement	Quantile forecasts improved accuracy and interpretation	H3, H4
Fan et al. (2023)	Model comparison	RMSE = 0.3108; R^2 = 0.4577	Preferred hybrid route led among reported routes	H3
Nazemi and Fabozzi (2024)	Interpretable ML	2,636 defaulted corporate bonds	Favorable out-of-time prediction and interpretable effects	H3, H4

Table 4 has summarized published evidence concerning validation reliability, forecasting robustness, and interpretability. Olson et al. (2021) have compared three cross-validation procedures and have demonstrated that preserving chronological information has contributed to more stable out-of-time forecasting. This result has aligned closely with Statistical Learning Theory because improved training performance has not necessarily guaranteed predictive generalization. Kellner et al. (2022) have combined quantile forecasting with feature interpretation and have reported precision improvements reaching 30% relative to their benchmark. Nazemi and Fabozzi (2024) have examined interpretable machine learning using 2,636 defaulted corporate bonds and have identified economically meaningful recovery relationships. Collectively, these studies have provided supportive evidence for H3 and H4. However, the findings have addressed two different characteristics of forecasting systems. Validation has concerned performance on observations beyond the estimation sample, whereas explainability has concerned understanding how model variables have influenced predictions. The evidence has therefore supported treating predictive reliability and transparency as related but distinct dimensions of model assessment.

Figure 8. Comparative Learning and Testing Evidence



Cross-case comparison of recovery predictions

Table 5: Recovery Prediction Across Financial Contexts

Study	Financial context	Numerical evidence	Comparative finding	Hypothesis
Olson et al. (2021)	U.S. corporate debt	12 machine learning methods	Validation choice influenced stability	H3, H5
Gambetti et al. (2022)	Corporate bonds	768 observations; mean recovery 30.98%	Security and economic predictors were relevant	H2, H5
Kellner et al. (2022)	U.S. and European bank loans	Macroeconomic importance up to 2× in U.S.	Regional predictor differences appeared	H5
Fan et al. (2023)	U.S. small-business banks	3 banks; 39,943 defaults	Hybrid models addressed heterogeneous losses	H1, H5
Nazemi and Fabozzi (2024)	U.S. corporate bond market	2,636 defaulted bonds	Interpretable models improved forecasting in setting	H4, H5

Table 5 has compared recovery prediction evidence across several financial environments and has addressed the third research question. The studies have represented corporate debt markets, corporate bonds, commercial banking portfolios, and small-business lending. Gambetti et al. (2022) have reported an average recovery rate of 30.98% across 768 corporate bond observations, while Fan et al. (2023) have examined 39,943 defaults across three American small-business banking datasets. These findings have not been interpreted as directly comparable recovery rates because the studies have employed different financial instruments and outcome measures. Kellner et al. (2022) have identified notable geographical differences in predictor importance, with macroeconomic information demonstrating substantially greater importance in American observations than in European observations. Statistical Learning Theory has provided a basis for interpreting such variation because models trained under one distribution of economic and institutional characteristics may not generalize equally across other environments. Consequently, the findings have suggested contextual differences relevant to H5, but they have not demonstrated that institutional variation independently causes changes in forecasting performance.

Appraisal of the included studies

The five research hypotheses have been evaluated using a qualitative five-point evidence assessment. The scale has represented the strength and direction of the available published evidence rather than responses collected from participants.

Table 6: Five-Point Evidence Assessment of Research Hypotheses

Hypothesis	Principal research variable	Evidence rating (1–5)	Evidence interpretation	Objective addressed
H1	Machine learning vs. conventional forecasting	3	Mixed; conditional advantages	Objective 1
H2	Recovery determinants and predictors	4	Supportive evidence	Objective 2
H3	Validation and generalization	4	Supportive evidence	Objective 3
H4	Explainability and interpretability	4	Supportive evidence	Objective 3
H5	Institutional/economic variation	3	Mixed; context differences identified	Objective 4

Table 6 has integrated the published evidence to evaluate the five hypotheses in relation to the research objectives. H1 has received a provisional rating of 3 because multiple studies have reported machine learning advantages, while differences in model specifications, datasets, and validation procedures have prevented a universal conclusion. H2 has received a rating of 4 because several studies have identified

relevant recovery determinants and improvements associated with informative predictors. H3 has received supportive assessment through published evidence emphasizing chronological validation and out-of-time forecasting. H4 has also received supportive assessment because interpretable modeling studies have demonstrated methods for explaining influential recovery characteristics. H5 has received a provisional rating of 3 because geographical and institutional differences have been documented without establishing a consistent quantitative relationship between those differences and model performance. Statistical Learning Theory has connected the five hypotheses through the concepts of informative learning, predictive error, model complexity, and generalization. These ratings have summarized the direction of the reviewed evidence and have not constituted statistical hypothesis tests, respondent-based Likert scores, or definitive confirmation of the hypotheses.

Discussion

Principal findings

The findings of this systematic literature review have demonstrated that machine learning approaches have exhibited considerable potential for improving nonperforming corporate loan recovery forecasting, although their predictive advantages have varied according to model architecture, recovery characteristics, and validation procedures. The comparative evidence presented in Sections 4.1–4.6 has provided conditional support for H1, stronger qualitative support for H2–H4, and mixed evidence concerning H5.

Comparison with previous evidence

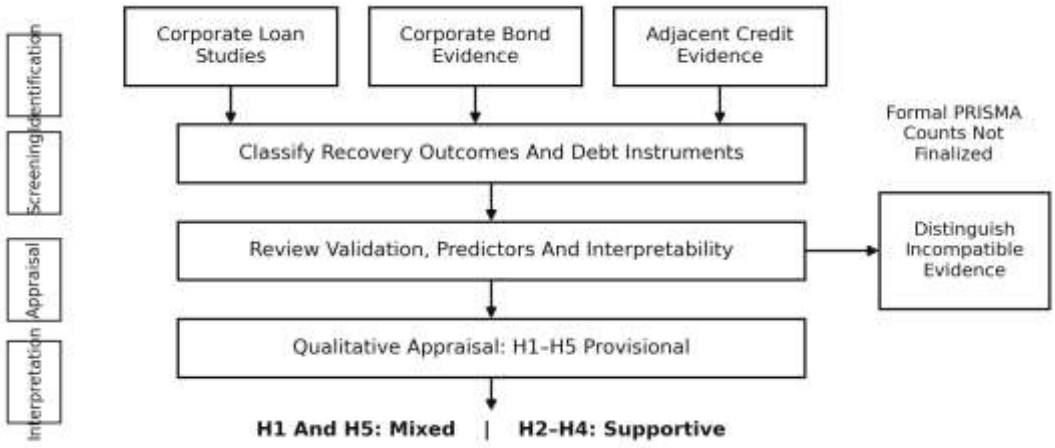
These findings have aligned with earlier investigations that identified nonlinear relationships and heterogeneous recovery outcomes as important considerations in credit loss prediction. In their comparative examination of 12 machine learning approaches, Olson et al. established that maintaining chronological order during cross-validation contributed to more stable out-of-time forecasts, reinforcing the present study's emphasis on generalization rather than training accuracy alone (Ashrafuzzaman, 2024; Olson et al., 2021). Similarly, research comparing 20 candidate algorithms found that combining forecasting methods with theoretically relevant financial information produced favorable corporate bond recovery predictions, supporting the importance of appropriate predictor selection (Gambetti et al., 2022; Jannatul, 2024). The present synthesis has also identified the relevance of previously reported numerical findings, including the MAE of 0.1835, RMSE of 0.3108, and R^2 of 0.4577 documented for a preferred hybrid model configuration using American small-business lending information (Fan et al., 2023; Kynatun & Saba, 2024). These values have represented individual published results rather than independently generated estimates or pooled statistical outcomes. Furthermore, evidence concerning quantile neural networks has demonstrated improvements reaching 30% in forecast precision relative to the original study's benchmark, particularly for higher loss quantiles (Kellner et al., 2022; Mashiat, 2024). Collectively, these comparisons have indicated that machine learning performance depends on the interaction of model selection, recovery information, and evaluation procedures. The findings have therefore addressed the principal research objective without establishing universal algorithmic superiority across corporate lending environments.

Mechanisms and implications for practice

The practical implications of the reviewed findings have centered on improving the reliability, transparency, and institutional applicability of corporate loan recovery forecasting within commercial banking operations. The evidence has indicated that banks cannot evaluate recovery models solely through predictive accuracy because collateral realization, workout expenses, borrower circumstances, and economic conditions have influenced actual recovery outcomes. The present synthesis has highlighted the importance of incorporating borrower-specific financial characteristics, loan seniority, collateral coverage, previous collection activity, and macroeconomic information into recovery prediction frameworks. Earlier research demonstrated that Cubist models, boosted trees, and random forests performed favorably in forecasting nonperforming retail loan collections, particularly when information concerning earlier recovery activities was included (Bellotti et al., 2021; Sumaiya, 2024). Although those findings have originated from retail exposures, they have provided methodological evidence concerning the predictive value of collection histories. Research on explainable credit-loss modeling has additionally demonstrated that complex algorithms can provide interpretable information concerning influential financial characteristics without necessarily sacrificing predictive accuracy (Bastos & Matos, 2022; Feroz, 2024). These findings have supported the practical

relevance of H2 and H4 by identifying opportunities for financial institutions to combine informative variables with transparent forecasting procedures. The reviewed evidence has also suggested that predicted recovery outcomes should be distinguished from realized collections when informing restructuring, collateral management, debt valuation, and portfolio-level loss assessment. An especially important interpretation has concerned the relationship between prediction accuracy and financial decision quality. A relatively accurate recovery forecast does not necessarily identify the most appropriate recovery intervention, particularly when legal expenses, repayment timing, and alternative settlement arrangements have differed. Consequently, the study has identified recovery prediction as one component of a broader institutional decision process requiring financial judgment, operational expertise, and documented model governance.

Figure 9. Interpretation of the Five Research Hypotheses



The theoretical interpretation of the findings has reinforced the relevance of Statistical Learning Theory as the principal foundation for understanding machine learning performance in nonperforming corporate loan recovery. Statistical Learning Theory has emphasized that successful predictive modeling depends on minimizing generalization error rather than reproducing patterns contained exclusively within historical training observations. The present review has incorporated this distinction by evaluating published evidence concerning predictive accuracy, model complexity, validation reliability, feature relevance, and institutional generalizability. The conceptual relationship between training error and out-of-sample predictive risk has provided a theoretical explanation for the mixed evidence associated with H1 and the stronger support assigned to H3. Earlier theoretical research established that model complexity and regularization influence predictive generalization, particularly when observations are limited or unevenly distributed (Cherkassky & Ma, 2009; Jahanara, 2025). The present findings have reflected this theoretical argument through the observation that increasingly sophisticated algorithms have not consistently demonstrated superior predictive reliability across different recovery environments. The study has also extended the interpretive application of Statistical Learning Theory by connecting predictor information, explanatory transparency, and recovery-specific institutional characteristics within a unified conceptual framework. Research examining interpretable machine learning for 2,636 defaulted American corporate bonds demonstrated that directly interpretable approaches could outperform several conventional and machine learning alternatives while revealing economically meaningful recovery relationships (Jannatul, 2025; Nazemi & Fabozzi, 2024). These findings have challenged the assumption that high predictive accuracy necessarily requires sacrificing model interpretability. Nevertheless, the review has distinguished explanatory consistency from causal interpretation because variables identified as important by learning algorithms have not necessarily represented independent causes of recovery outcomes. The theoretical contribution has therefore consisted of organizing published evidence around predictive generalization,

information relevance, interpretability, and contextual heterogeneity without claiming that these relationships have been independently estimated or statistically confirmed.

Based on the synthesized findings, financial institutions should prioritize transparent, robust, and context-sensitive machine learning approaches for forecasting nonperforming corporate loan recovery. Recovery prediction frameworks should incorporate borrower characteristics, collateral information, contractual conditions, historical collections, and macroeconomic indicators. Banks should implement rigorous out-of-time validation, standardized performance evaluation, and explainable artificial intelligence techniques to strengthen forecasting reliability. Furthermore, researchers should investigate the proposed Temporal, Explainable, and Recovery-Aware Machine Learning Framework (TERM-LR), integrating ensemble algorithms, survival analysis, and uncertainty estimation. Comparative investigations across banking institutions and economic environments should employ consistent recovery definitions, independent validation procedures, and reliable corporate lending information.

Limitations

Several methodological limitations have affected the interpretation of the findings and have required careful consideration when comparing the reviewed studies. First, the available literature has included corporate bank loans, publicly traded corporate bonds, small-business credit, leasing agreements, and retail nonperforming exposures, each representing different contractual structures, recovery procedures, and outcome definitions. These differences have restricted the direct comparability of published predictive-performance measures. Second, the review has identified substantial variation in training datasets, forecast horizons, explanatory variables, validation strategies, and reporting standards. Such heterogeneity has prevented the defensible calculation of a common pooled performance estimate from the evidence presented. Third, the present research has employed qualitative synthesis rather than independently trained forecasting models, and its five-point evidence ratings have represented provisional interpretive judgments rather than respondent-based Likert observations or statistical significance tests. Furthermore, a complete PRISMA screening record and independently verified quality appraisal have not yet been established, so the reported evidence cannot be treated as an exhaustive systematic review sample. Another important limitation has concerned the incomplete resolution of defaulted corporate loans. Recovery cases that remain under negotiation or liquidation can produce censored observations, potentially affecting recovery estimates when only completed cases are analyzed. Research examining corporate debt recovery and time to resolution has demonstrated the importance of jointly considering recovery severity and resolution duration when addressing such estimation problems (Jacobs, 2024; Mostari, 2025). Differences between American and European banking environments have also been documented, with macroeconomic characteristics demonstrating substantially different predictive importance across the examined regions (Kellner et al., 2022; Mashfiqur, 2025). These limitations have provided a basis for interpreting H5 cautiously and have reinforced the distinction between published associations and generalizable predictive relationships.

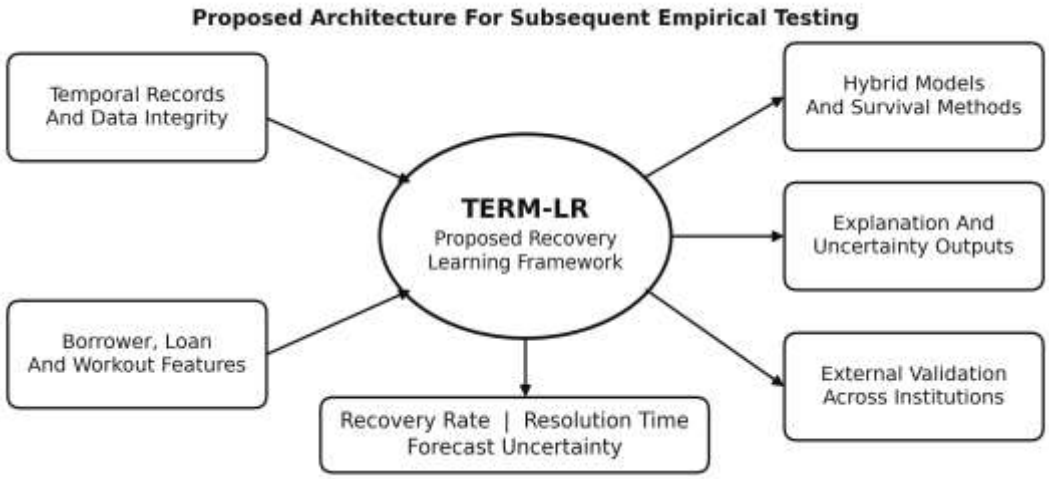
This research has encountered several limitations associated with its qualitative systematic literature review design and dependence on previously published evidence. Differences in financial instruments, borrower characteristics, recovery definitions, institutional environments, and predictive evaluation procedures have restricted meaningful comparisons across studies. The absence of primary data collection, independently trained models, and quantitative meta-analysis has prevented original statistical estimation and definitive hypothesis confirmation. Furthermore, incomplete reporting of forecasting procedures, potential publication bias, and limited availability of corporate loan recovery datasets have constrained evidence interpretation. The provisional five-point evidence assessments have remained qualitative judgments, while the absence of completed PRISMA screening and comprehensive quality appraisal has limited systematic review completeness.

Future research

Future research should prioritize the development and empirical validation of an integrated Temporal, Explainable, and Recovery-Aware Machine Learning Framework (TERM-LR) specifically designed for nonperforming corporate loan recovery forecasting. The proposed framework would address the methodological weaknesses identified in the present review through five interconnected components:

temporally structured data processing, recovery-specific feature integration, hybrid predictive modeling, explainable forecasting, and cross-institutional validation. First, researchers could construct longitudinal corporate loan recovery datasets containing borrower financial statements, collateral valuations, loan seniority, repayment histories, restructuring arrangements, insolvency proceedings, recovery costs, and macroeconomic indicators. Second, the framework could integrate gradient boosting and random forests with survival analysis and quantile prediction techniques to estimate not only the expected recovery proportion but also recovery timing, uncertainty, and the probability of alternative resolution outcomes. Its predictive architecture could jointly examine discounted recovery rates and time to resolution while distinguishing complete recoveries, partial recoveries, and unresolved defaults.

Figure 10. Proposed TERM-LR Research Framework



Third, explainable artificial intelligence methods, including SHAP-based feature attribution, accumulated local effects, and interpretable generalized additive models, could provide transparent representations of the factors associated with individual recovery estimates. Fourth, researchers could incorporate institution-aware learning or domain adaptation techniques to investigate whether forecasting models trained on one banking environment maintain their performance across different jurisdictions, corporate sectors, and economic conditions. Finally, the framework should be evaluated through sequential temporal validation, institution-held-out testing, MAE, RMSE, quantile loss, uncertainty calibration, and comparisons against conventional regression and established machine learning benchmarks. This proposed architecture would extend the multistage modeling principles previously investigated for bank loan losses while incorporating the temporal stability concerns identified in corporate debt forecasting research (Fan et al., 2023; Mashiat, 2025; Olson et al., 2021). It would also build upon evidence that transparent recovery models can reveal financially meaningful predictor relationships (Arif Uz et al., 2025; Nazemi & Fabozzi, 2024). An additional research direction would involve assessing recovery decision quality by comparing the economic consequences of model-assisted restructuring, collateral enforcement, and settlement strategies rather than evaluating forecasting error alone. Such investigations would require actual institutional records, carefully defined decision objectives, and appropriate controls for selection bias. TERM-LR remains a proposed conceptual model, not an empirically developed or validated result of the present study. Its central research contribution would be to establish a testable framework connecting predictive performance, recovery timing, explainability, institutional adaptability, and financially meaningful outcomes within corporate nonperforming loan management.

Conclusion

This research has systematically examined published evidence concerning machine learning applications in nonperforming corporate loan recovery, emphasizing predictive performance, recovery determinants,

model reliability, explainability, and institutional differences. The findings have indicated that machine learning approaches, particularly ensemble methods, tree-based algorithms, and neural networks, have demonstrated favorable predictive capabilities under specific conditions. Borrower characteristics, collateral arrangements, contractual features, and macroeconomic indicators have emerged as important recovery predictors. Statistical Learning Theory has provided a theoretical foundation for interpreting model generalization and predictive reliability. The five research hypotheses have received varying degrees of qualitative support, although methodological heterogeneity has prevented definitive statistical confirmation or pooled performance estimation.

REFERENCES

[1]. Abdullah Mohammad, S. (2023). $CUGAO_2$ and $CUINS_2$ Quantum Dot Sensitization for Enhanced Photovoltaic Efficiency in Wearable Fiber-Shaped Solar Cells. *American Journal of Advanced Technology and Engineering Solutions*, *3*(04), 209-261. https://doi.org/10.63125/wysdvy71

[2]. Abdullah Mohammad, S. (2024). Triboelectric And Piezoelectric Nanogenerators Using Binary and Ternary Oxide Nanostructures: Optimization of Energy Conversion Efficiency for Wearable Harvesting. *Review of Applied Science and Technology*, *3*(03), 31–72. https://doi.org/10.63125/w7q9cs55

[3]. Acharya, V. V., Bharath, S. T., & Srinivasan, A. (2007). Does industry-wide distress affect defaulted firms? Evidence from creditor recoveries. *Journal of Financial Economics*, *85*(3), 787–821. https://doi.org/10.1016/j.jfineco.2006.05.011

[4]. Adnan, S., & Iftekhar, A. (2023). Predicting Cost Variance in Digital Transformation Projects Through Ensemble Learning and Project Performance Analytics. *International Journal of Scientific Interdisciplinary Research*, *13*(10), 01–36. https://doi.org/10.63125/j18hxs51

[5]. Ahmed, H. C. H. R. (2024). Machine Learning Enhanced Demand Forecasting and Portfolio Analytics for Scalable Growth Management in Digital Logistics Operations. *American Journal of Advanced Technology and Engineering Solutions*, *4*(02), 33-81. https://doi.org/10.63125/5xq3nd53

[6]. Ahmed, H. C. H. R. (2026). An Artificial Intelligence Integrated IT Management Framework for Analytics-Driven Operational Decision-Making in U.S. Supply Chain and Logistics Systems. *American Journal of Interdisciplinary Studies*, *16*(06), 01-37. https://doi.org/10.63125/b5hgg259

[7]. Ahmed, H. C. H. R., & Nuzhat Sadia, P. (2023). A Quantitative Assessment of CRM Analytics and KPI Dashboard Design on Enterprise Revenue Conversion in Emerging Market Logistics Platforms. *American Journal of Data Science and Analytics*, *4*(09), 50-95. https://doi.org/10.63125/315bge15

[8]. Altman, E. I., Brady, B., Resti, A., & Sironi, A. (2005). The link between default and recovery rates: Theory, empirical evidence, and implications. *The Journal of Business*, *78*(6), 2203–2228. https://doi.org/10.1086/497044

[9]. Altman, E. I., & Kalotay, E. A. (2014). Ultimate recovery mixtures. *Journal of Banking & Finance*, *40*, 116–129. https://doi.org/10.1016/j.jbankfin.2013.11.021

[10]. Aminul Islam, M. (2024). Interpreting Cyber-Threat Indicators in Enterprise Memory Forensics: A Mixed-Methods Study of Machine Learning Performance and Explainability. *Review of Applied Science and Technology*, *14*(12), 01–13. https://doi.org/10.63125/2p8hte92

[11]. Aminul Islam, M., & Md, A. (2023). Adaptive Class-Imbalance Learning for High-Sensitivity Fraud Detection in Large-Scale Digital Transaction Networks. *International Journal of Business and Economics Insights*, *13*(09), 01-33. https://doi.org/10.63125/bmxwv741

[12]. Asheke, M. (2023). Machine Learning–Based Predictive Quality Modeling for Defect Detection and Process Optimization in Advanced Manufacturing Systems. *Review of Applied Science and Technology*, *13*(12), 01–49. https://doi.org/10.63125/thy73b16

[13]. Asheke, M. (2024). Digital Twin and Industrial Internet of Things–Enabled Predictive Maintenance for Reliability Enhancement in Smart Manufacturing Systems. *International Journal of Scientific Interdisciplinary Research*, *5*(2), 792–838. https://doi.org/10.63125/jgrdqd62

[14]. Ashik, M., & Shabbir, A. (2022). Digital Twin-Enabled Predictive Maintenance for Rotating and Static Equipment: Bridging FEA-Based Failure Analysis with Real-Time Asset Health Monitoring in Energy and Power Facilities. *International Journal of Scientific Interdisciplinary Research*, *12*(05), 01–37. https://doi.org/10.63125/0qaysz49

[15]. Ashrafuzzaman, M. (2024). The impact of cloud-based management information systems on hrm efficiency: an analysis of small and medium-sized enterprises (SMES). *Academic Journal on Artificial Intelligence, Machine Learning, Data Science and Management Information Systems*, *1*(01), 40-56. https://doi.org/10.69593/ajaimldsmis.v1i01.124

[16]. Bandyopadhyay, A. (2022). Loan level loss given default (LGD) study of Indian banks. *IIMB Management Review*, *34*(2), 168–177. https://doi.org/10.1016/j.iimb.2022.06.003

[17]. Bastos, J. A. (2010). Forecasting bank loans loss-given-default. *Journal of Banking & Finance*, *34*(10), 2510–2517. https://doi.org/10.1016/j.jbankfin.2010.04.011

[18]. Bastos, J. A. (2014). Ensemble predictions of recovery rates. *Journal of Financial Services Research*, *46*(2), 177–193. https://doi.org/10.1007/s10693-013-0165-3

[19]. Bastos, J. A., & Matos, S. M. (2022). Explainable models of credit losses. *European Journal of Operational Research*, *301*(1), 386–394. https://doi.org/10.1016/j.ejor.2021.11.009

[20]. Bellotti, A., Brigo, D., Gambetti, P., & Vrins, F. (2021). Forecasting recovery rates on non-performing loans with machine learning. *International Journal of Forecasting*, *37*(1), 428–444. https://doi.org/10.1016/j.ijforecast.2020.06.009

[21]. Betz, J., Kellner, R., & Rösch, D. (2018). Systematic effects among loss given defaults and their implications on downturn estimation. *European Journal of Operational Research*, *271*(3), 1113–1144. https://doi.org/10.1016/j.ejor.2018.05.059

[22]. Bonfim, D., Dias, D. A., & Richmond, C. (2012). What happens after corporate default? Stylized facts on access to credit. *Journal of Banking & Finance*, *36*(7), 2007–2025. https://doi.org/10.1016/j.jbankfin.2012.03.007

[23]. Bruche, M., & González-Aguado, C. (2010). Recovery rates, default probabilities, and the credit cycle. *Journal of Banking & Finance*, *34*(4), 754–764. https://doi.org/10.1016/j.jbankfin.2009.04.009

[24]. Calabrese, R. (2014). Predicting bank loan recovery rates with a mixed continuous-discrete model. *Applied Stochastic Models in Business and Industry*, *30*(2), 99–114. https://doi.org/10.1002/asmb.1932

[25]. Calabrese, R., & Zenga, M. (2010). Bank loan recovery rates: Measuring and nonparametric density estimation. *Journal of Banking & Finance*, *34*(5), 903–911. https://doi.org/10.1016/j.jbankfin.2009.10.001

[26]. Carannante, M., D'Amato, V., Fersini, P., Forte, S., & Melisi, G. (2024). Machine learning due diligence evaluation to increase NPLs profitability transactions on secondary market. *Review of Managerial Science*, *18*(7), 1963–1983. https://doi.org/10.1007/s11846-023-00635-y

[27]. Caselli, S., Gatti, S., & Querci, F. (2008). The sensitivity of the loss given default rate to systematic risk: New empirical evidence on bank loans. *Journal of Financial Services Research*, *34*(1), 1–34. https://doi.org/10.1007/s10693-008-0033-8

[28]. Chen, X., Wang, G., & Zhang, X. (2019). Modeling recovery rate for leveraged loans. *Economic Modelling*, *81*, 231–241. https://doi.org/10.1016/j.econmod.2019.04.006

[29]. Cherkassky, V., & Ma, Y. (2009). Another look at statistical learning theory and regularization. *Neural Networks*, *22*(7), 958–969. https://doi.org/10.1016/j.neunet.2009.04.005

[30]. Chuang, H.-C., & Chen, J.-E. (2023). Exploring industry-distress effects on loan recovery: A double machine learning approach for quantiles. *Econometrics*, *11*(1), 6. https://doi.org/10.3390/econometrics11010006

[31]. Couwenberg, O., & de Jong, A. (2008). Costs and recovery rates in the Dutch liquidation-based bankruptcy system. *European Journal of Law and Economics*, *26*(2), 105–127. https://doi.org/10.1007/s10657-008-9058-6

[32]. Davydenko, S. A., & Franks, J. R. (2008). Do bankruptcy codes matter? A study of defaults in France, Germany, and the U.K. *The Journal of Finance*, *63*(2), 565–608. https://doi.org/10.1111/j.1540-6261.2008.01325.x

[33]. Dermine, J., & Neto de Carvalho, C. (2006). Bank loan losses-given-default: A case study. *Journal of Banking & Finance*, *30*(4), 1219–1243. https://doi.org/10.1016/j.jbankfin.2005.05.005

[34]. Fan, M., Wu, T.-H., & Zhao, Q. (2023). Assessing the loss given default of bank loans using the hybrid algorithms multi-stage model. *Systems*, *11*(10), 505. https://doi.org/10.3390/systems11100505

[35]. Fenech, J. P., Yap, Y. K., & Shafik, S. (2016). Modelling the recovery outcomes for defaulted loans: A survival analysis approach. *Economics Letters*, *145*, 79–82. https://doi.org/10.1016/j.econlet.2016.05.015

[36]. Gambetti, P., Gauthier, G., & Vrins, F. (2019). Recovery rates: Uncertainty certainly matters. *Journal of Banking & Finance*, *106*, 371–383. https://doi.org/10.1016/j.jbankfin.2019.07.010

[37]. Gambetti, P., Roccazzella, F., & Vrins, F. (2022). Meta-learning approaches for recovery rate prediction. *Risks*, *10*(6), 124. https://doi.org/10.3390/risks10060124

[38]. Gazi Yeasin, A. (2023). Impact of Accessibility-First Design Systems on WCAG 2.1 Compliance and Inclusive User Experience in Enterprise Software Platforms: A Mixed-Methods Evaluation. *Review of Applied Science and Technology*, *13*(07), 01–49. https://doi.org/10.63125/nxh4gp19

[39]. Gostkowski, M., & Gajowniczek, K. (2020). Weighted quantile regression forests for bimodal distribution modeling: A loss given default case. *Entropy*, *22*(5), 545. https://doi.org/10.3390/e22050545

[40]. Grunert, J., & Weber, M. (2009). Recovery rates of commercial lending: Empirical evidence for German companies. *Journal of Banking & Finance*, *33*(3), 505–513. https://doi.org/10.1016/j.jbankfin.2008.09.002

[41]. Gupton, G. M. (2005). Advancing loss given default prediction models: How the quiet have quickened. *Economic Notes*, *34*(2), 185–230. https://doi.org/10.1111/j.0391-5026.2005.00149.x

[42]. Gürtler, M., & Hibbeln, M. (2013). Improvements in loss given default forecasts for bank loans. *Journal of Banking & Finance*, *37*(7), 2354–2366. https://doi.org/10.1016/j.jbankfin.2013.01.031

[43]. Gürtler, M., & Zöllner, M. (2023). Heterogeneities among credit risk parameter distributions: The modality defines the best estimation method. *OR Spectrum*, *45*(1), 251–287. https://doi.org/10.1007/s00291-022-00689-6

[44]. Han, C., & Jang, Y. (2013). Effects of debt collection practices on loss given default. *Journal of Banking & Finance*, *37*(1), 21–31. https://doi.org/10.1016/j.jbankfin.2012.08.009

[45]. Hartmann-Wendels, T., Miller, P., & Töws, E. (2014). Loss given default for leasing: Parametric and nonparametric estimations. *Journal of Banking & Finance*, *40*, 364–375. https://doi.org/10.1016/j.jbankfin.2013.12.006

[46]. Höcht, S., Min, A., Wieczorek, J., & Zagst, R. (2022). Explaining aggregated recovery rates. *Risks*, *10*(1), 18. https://doi.org/10.3390/risks10010018

[47]. Hurlin, C., Leymarie, J., & Patin, A. (2018). Loss functions for loss given default model comparison. *European Journal of Operational Research*, *268*(1), 348–360. https://doi.org/10.1016/j.ejor.2018.01.020

[48]. Ingermann, P.-H., Hesse, F., Bélorgey, C., & Pfingsten, A. (2016). The recovery rate for retail and commercial customers in Germany: A look at collateral and its adjusted market values. *Business Research*, *9*(2), 179–228. https://doi.org/10.1007/s40685-016-0028-5

[49]. Jaber, J. J., Ismail, N., Ramli, S. N. M., Albadareen, B., & Hamadneh, N. N. (2021). Estimating loss given default based on beta regression. *Computers, Materials & Continua*, *66*(3), 3329–3344. https://doi.org/10.32604/cmc.2021.014509

[50]. Jacobs, M., Jr. (2024). Modeling ultimate loss-given-default and time-to-resolution on corporate debt. *Journal of Financial Risk Management*, *13*(2), 426–459. https://doi.org/10.4236/jfrm.2024.132020

[51]. Jahanara, A. (2025). AI-Enabled Workforce Management: Its Impact on Employee Burnout, Staffing Efficiency, and Organizational Performance in Healthcare. *Review of Applied Science and Technology*, *4*(04), 247–295. https://doi.org/10.63125/5617ff74

[52]. Jahanara, A. (2026). Impact of Artificial Intelligence on Operational Efficiency and Cost Reduction in U.S. Healthcare Organizations. *American Journal of Interdisciplinary Studies*, *7*(02), 225-276. https://doi.org/10.63125/3p24ma43

[53]. Jahanara, A., & Mst Kaniz, F. (2023). Demand Variability and Capacity Planning in Home and Community-Based Services: A Decision-Analytic Framework for Health Services Managers. *American Journal of Data Science and Analytics*, *4*(09), 01-49. https://doi.org/10.63125/3y62vr16

[54]. Jankowitsch, R., Nagler, F., & Subrahmanyam, M. G. (2014). The determinants of recovery rates in the US corporate bond market. *Journal of Financial Economics*, *114*(1), 155–177. https://doi.org/10.1016/j.jfineco.2014.06.001

[55]. Jannatul, F. (2024). Alzheimer's Disease and Mobility Multimodal Prediction of Gait Decline, Fall Risk, and Loss of Functional Independence in Mild Cognitive Impairment and Early Alzheimer's Disease. *Review of Applied Science and Technology*, *3*(01), 420–483. https://doi.org/10.63125/q2d90m63

[56]. Jannatul, F. (2025). Spinal Cord Injury and Neurorehabilitation Deep Learning–Based Prediction of Functional Recovery Trajectories After Traumatic Spinal Cord Injury. *ASRC Procedia: Global Perspectives in Science and Scholarship*, *1*(01), 2865–2922. https://doi.org/10.63125/m738bh82

[57]. Jannatul, F. (2026). Predicting Falls and Mobility Decline Among Older Adults Using Longitudinal Functional Data and Ensemble Machine Learning. *International Journal of Scientific Interdisciplinary Research*, *7*(1), 581–631. https://doi.org/10.63125/56hans40

[58]. Kaposty, F., Kriebel, J., & Löderbusch, M. (2020). Predicting loss given default in leasing: A closer look at models and variable selection. *International Journal of Forecasting*, *36*(2), 248–266. https://doi.org/10.1016/j.ijforecast.2019.05.009

[59]. Kazi Rakib Hasan, S., & Adnan, S. (2022). Quantitative Assessment of Machine Learning-Based Schedule Risk Prediction for Reducing Delays in Complex Technology Projects. *Review of Applied Science and Technology*, *1*(04), 464–501. https://doi.org/10.63125/dn8nkm85

[60]. Kellner, R., Nagl, M., & Rösch, D. (2022). Opening the black box: Quantile neural networks for loss given default prediction. *Journal of Banking & Finance*, *134*, 106334. https://doi.org/10.1016/j.jbankfin.2021.106334

[61]. Khieu, H. D., Mullineaux, D. J., & Yi, H.-C. (2012). The determinants of bank loan recovery rates. *Journal of Banking & Finance*, *36*(4), 923–933. https://doi.org/10.1016/j.jbankfin.2011.10.005

[62]. Krüger, S., & Rösch, D. (2017). Downturn LGD modeling using quantile regression. *Journal of Banking & Finance*, *79*, 42–56. https://doi.org/10.1016/j.jbankfin.2017.03.001

[63]. Kynatun, M., & Saba, A. (2024). Machine Learning Enhanced Intrusion Detection and Anomaly Classification for MIS-Driven Enterprise Network Security. *International Journal of Scientific Interdisciplinary Research*, *14*(09), 01–41. https://doi.org/10.63125/xk6kpc72

[64]. Loterman, G., Brown, I., Martens, D., Mues, C., & Baesens, B. (2012). Benchmarking regression algorithms for loss given default modeling. *International Journal of Forecasting*, *28*(1), 161–170. https://doi.org/10.1016/j.ijforecast.2011.01.006

[65]. Maishar Mostari, Z. (2025). Impact of Deep Learning on Real-Time Network Intrusion Detection: A Quantitative Evaluation Using Benchmark Datasets. *Review of Applied Science and Technology*, *4*(04), 296–338. https://doi.org/10.63125/tnaksd40

[66]. Mansura Akter, E., & Sumaiya, M. (2022). Antimicrobial Resistance Profiles of Acinetobacter Baumannii and Vibrio Cholerae in Hospital Wastewater and Adjacent Household Water Sources. *American Journal of Data Science and Analytics*, *3*(09), 01-55. https://doi.org/10.63125/gbervt14

[67]. Mashfiqur, R. (2025). Multivariate Time-Series Forecasting of Working Capital Requirements for Improving Cash-Flow Stability in Export-Oriented Manufacturing. *American Journal of Data Science and Analytics*, *6*(12), 217-249. https://doi.org/10.63125/ty3dtt46

[68]. Mashfiqur, R. (2026). Explainable Artificial Intelligence for Supplier Cost-Risk Assessment and Margin Optimization in Global Manufacturing Supply Chains. *American Journal of Interdisciplinary Studies*, *16*(08), 01-37. https://doi.org/10.63125/np6bw920

[69]. Mashfiqur, R., & Md, A. (2022). Monte Carlo Simulation and Stochastic Cash-Flow Modeling for Reducing Liquidity Risk in Project-Based Enterprises. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 382–431. https://doi.org/10.63125/zf3fwv21

[70]. Mashiat, N. (2024). Reducing Telecommunications Service Interruptions Through Predictive Root-Cause Analytics of Multivariate System Telemetry. *International Journal of Scientific Interdisciplinary Research*, *14*(11), 01-37. https://doi.org/10.63125/nqjh0c29

[71]. Mashiat, N. (2025). Improving Enterprise Data Integrity Through Automated Validation and Schema-Drift Detection in Large-Scale Data Integration Systems. *Review of Applied Science and Technology*, *4*(02), 889–934. https://doi.org/10.63125/21zzkk04

[72]. Matenda, F. R., Sibanda, M., Chikodza, E., & Gumbo, V. (2022). Corporate loan recovery rates under downturn conditions in a developing economy: Evidence from Zimbabwe. *Risks*, *10*(10), 198. https://doi.org/10.3390/risks10100198

[73]. Md, A., & Gazi Yeasin, A. (2022). Machine Learning Enhanced Spatial-Data Workflow Redesign: A Quantitative Evaluation of Task Success and Operational Cost Reduction in Enterprise LiDAR Platforms. *American Journal of Data Science and Analytics*, *3*(06), 80-136. https://doi.org/10.63125/0den7c02

[74]. Md Arif Uz, Z., Sharmin, S., & Md Abdur, R. (2025). Factors Influencing Behavioural Intention to Use he Innovative Open And Distance Learning Systems With The Role Of Continuance Intention. *American Journal of Scholarly Research and Innovation*, *4*(01), 202-220. https://doi.org/10.63125/pbjxp014

[75]. Md Ashfaq, S., & Abdullah Mohammad, S. (2022). CNT/PANI Composite Electrodes for Flexible Fiber-Shaped Dye-Sensitized Solar Cells: Synthesis, Characterization, and Photovoltaic Performance. *American Journal of Interdisciplinary Studies*, *3*(02), 221-274. https://doi.org/10.63125/m334mf41

[76]. Md Asif Ali Sheak, A., & Tahmedur, R. (2023). A Data-Driven Framework for Engineering Project Risk Assessment Under Schedule and Cost Uncertainty. *International Journal of Scientific Interdisciplinary Research*, *4*(2), 204–251. https://doi.org/10.63125/kq4zzh16

[77]. Md Imran, H., & Jannatul, F. (2021). Translational Stroke and Neurorehabilitation Research Integrating Neuroimaging, Circulating Biomarkers, and Functional Assessments to Predict Recovery After Stroke. *American Journal of Data Science and Analytics*, *2*(12), 156-198. https://doi.org/10.63125/c28rzv23

[78]. Md Imran, H., & Jannatul, F. (2022). Molecular and Precision Rehabilitation Molecular Mechanisms Linking Mitochondrial Dysfunction, Inflammation, and Muscle Degeneration with Functional Decline and Rehabilitation Response in Neuromuscular Disease. *American Journal of Interdisciplinary Studies*, *12*(08), 01-44. https://doi.org/10.63125/0ym5bh79

[79]. Md Kaium, H. (2023). A GIS-Based Multi-Criteria Decision-Support and Land-Suitability Optimization Framework for Climate-Resilient Urban and Regional Planning. *American Journal of Data Science and Analytics*, *4*(06), 43–103. https://doi.org/10.63125/wvq0sx87

[80]. Md Mahamud Pervaz, P. (2023). Integration of MES, SAP, And PI Historian Platforms for Real-Time Industrial Asset Performance Management. *Review of Applied Science and Technology*, *2*(04), 393–445. https://doi.org/10.63125/9ebtas97

[81]. Md Mahamud Pervaz, P. (2025). Development of a CMMS-Integrated Predictive Maintenance Framework for Industrial PLC and Drive Systems. *American Journal of Scholarly Research and Innovation*, *4*(01), 894–942. https://doi.org/10.63125/j33gtf31

[82]. Md. Morshedul, I., & Aminul Islam, M. (2022). A Quantitative Evaluation of Explainable Ensemble Learning for Probabilistic Risk Estimation and Decision Reliability in Enterprise Analytics. *American Journal of Data Science and Analytics*, *3*(02), 01-38. https://doi.org/10.63125/k22shk74

[83]. Mohammad Badrul, A., & Md. Mominul, H. (2023). Machine Learning–Based Assessment of Environmental and Public Health Risks in Sewerage Sludge Management Systems. *American Journal of Advanced Technology and Engineering Solutions*, *3*(02), 33-77. https://doi.org/10.63125/y3yfxa92

[84]. Mst Kaniz, F., & Jahanara, A. (2022). Lean Six Sigma Implementation in Home Health Administrative Operations: A DMAIC Case Study of Managerial Workflow Redesign and Error Reduction. *International Journal of Scientific Interdisciplinary Research*, *12*(04), 01–60. https://doi.org/10.63125/xchca204

[85]. Nagl, M., Nagl, M., & Rösch, D. (2022). Quantifying uncertainty of machine learning methods for loss given default. *Frontiers in Applied Mathematics and Statistics*, *8*, 1076083. https://doi.org/10.3389/fams.2022.1076083

[86]. Nazemi, A., Baumann, F., & Fabozzi, F. J. (2022). Intertemporal defaulted bond recoveries prediction via machine learning. *European Journal of Operational Research*, *297*(3), 1162–1177. https://doi.org/10.1016/j.ejor.2021.06.047

[87]. Nazemi, A., & Fabozzi, F. J. (2018). Macroeconomic variable selection for creditor recovery rates. *Journal of Banking & Finance*, *89*, 14–25. https://doi.org/10.1016/j.jbankfin.2018.01.006

[88]. Nazemi, A., & Fabozzi, F. J. (2024). Interpretable machine learning for creditor recovery rates. *Journal of Banking & Finance*, *164*, 107187. https://doi.org/10.1016/j.jbankfin.2024.107187

[89]. Nazemi, A., Fatemi Pour, F., Heidenreich, K., & Fabozzi, F. J. (2017). Fuzzy decision fusion approach for loss-given-default modeling. *European Journal of Operational Research*, *262*(2), 780–791. https://doi.org/10.1016/j.ejor.2017.04.008

[90]. Nazemi, A., Heidenreich, K., & Fabozzi, F. J. (2018). Improving corporate bond recovery rate prediction using multi-factor support vector regressions. *European Journal of Operational Research*, *271*(2), 664–675. https://doi.org/10.1016/j.ejor.2018.05.024

[91]. Nazemi, A., Rezazadeh, H., Fabozzi, F. J., & Höchstötter, M. (2022). Deep learning for modeling the collection rate for third-party buyers. *International Journal of Forecasting*, *38*(1), 240–252. https://doi.org/10.1016/j.ijforecast.2021.03.013

[92]. Nehrebecka, N. (2019). Bank loans recovery rate in commercial banks: A case study of non-financial corporations. *Zbornik radova Ekonomskog fakulteta u Rijeci / Proceedings of Rijeka Faculty of Economics: Journal of Economics and Business*, *37*(1), 139–172. https://doi.org/10.18045/zbefri.2019.1.139

[93]. Nishat, J. (2023). Communicable Disease Prevention in a Post-Pandemic Era: A Global Review of Public Health Education, Surveillance, and Community-Level Interventions. *International Journal of Scientific Interdisciplinary Research*, *4*(4), 630–678. https://doi.org/10.63125/mnqefj36

[94]. Nuzhat Sadia, P. (2023). A Multi-Criteria Decision-Support Framework For ERP-Based Business Process Optimization in Mid-Sized Manufacturing Enterprises. *American Journal of Scholarly Research and Innovation*, *2*(02), 519–573. https://doi.org/10.63125/j58k5d55

[95]. Olson, L. M., Qi, M., Zhang, X., & Zhao, X. (2021). Machine learning loss given default for corporate debt. *Journal of Empirical Finance*, *64*, 144–159. https://doi.org/10.1016/j.jempfin.2021.08.009

[96]. Qi, M., & Zhao, X. (2011). Comparison of modeling methods for loss given default. *Journal of Banking & Finance*, *35*(11), 2842–2855. https://doi.org/10.1016/j.jbankfin.2011.03.011

[97]. Rony, S., & Md, A. (2026). Multi-Agency Coordination in Sustained Healthcare Supply Shortages: Lessons from Covid-19, Drug Shortages, and Critical Medical Supply Disruptions. *American Journal of Data Science and Analytics*, *7*(01), 181-221. https://doi.org/10.63125/5v6t8515

[98]. Sadia, Z., & Maishar Mostari, Z. (2022). The Impact of Interactive Dashboard Visualization on Data-Driven Decision Making in Financial Services: A Qualitative Evaluation. *International Journal of Scientific Interdisciplinary Research*, *12*(03), 01–35. https://doi.org/10.63125/3jstsv75

[99]. Sopitpongstorn, N., Silvapulle, P., Gao, J., & Fenech, J.-P. (2021). Local logit regression for loan recovery rate. *Journal of Banking & Finance*, *126*, 106093. https://doi.org/10.1016/j.jbankfin.2021.106093

[100]. Stabak, D., & Md Shahid, N. (2022). Empirical Assessment of Uncertainty-Aware Real-Time Decision-Making for Reliable and Resilient Cyber-Physical Systems. *International Journal of Scientific Interdisciplinary Research*, *1*(01), 432–477. https://doi.org/10.63125/q3r4kk16

[101]. Stabak, D., & Md Siam, T. (2023). Real-Time Edge Computing for Cyber-Physical Systems: A Meta-Analysis of Energy Efficiency and Fault Tolerance. *American Journal of Scholarly Research and Innovation*, *13*(06), 01–44. https://doi.org/10.63125/g9c9cp41

[102]. Starosta, W. (2021). Loss given default decomposition using mixture distributions of in-default events. *European Journal of Operational Research*, *292*(3), 1187–1199. https://doi.org/10.1016/j.ejor.2020.11.034

[103]. Sumaiya, M. (2024). Hospital Wastewater as A Reservoir of Antimicrobial Resistance Genes: Implications for Public Health and Genomic Surveillance in The United States. *Review of Applied Science and Technology*, *3*(04), 448–508. https://doi.org/10.63125/ahvas976

[104]. Sumaiya, M. (2025). CRISPR-Cas Systems in Precision Oncology: Current Applications and Future Challenges. *American Journal of Interdisciplinary Studies*, *6*(1), 513-564. https://doi.org/10.63125/rqpqr418

[105]. Sumaiya, M. (2026). Artificial Intelligence Driven Multi-Omics Approaches for Precision Medicine: From Molecular Discovery to Clinical Translation. *American Journal of Scholarly Research and Innovation*, *5*(01), 429–477. https://doi.org/10.63125/yfme0q12

[106]. Syed Feroz, A. (2023). Quality-Management Systems for Resilient Energy Infrastructure Fabrication. *International Journal of Scientific Interdisciplinary Research*, *13*(10), 37–91. https://doi.org/10.63125/mbhhrm88

[107]. Syed Feroz, A. (2024). AI-Assisted NDT Analytics for Early Crack Detection in Energy-Infrastructure Pressure Vessels Using Advanced Manufacturing Data. *Review of Applied Science and Technology*, *14*(07), 01-52. https://doi.org/10.63125/8bdqs538

[108]. Syed Feroz, A. (2025). Intelligent Process Monitoring for Enhancing Weld Integrity and Reliability in Advanced Manufacturing of LPG Energy Infrastructure. *American Journal of Interdisciplinary Studies*, *6*(3), 394-441. https://doi.org/10.63125/p3b5mb44

[109]. Syed Feroz, A., & Md Mahamud Pervaz, P. (2022). Root-Cause Analysis and FMEA-Driven Reliability Improvement in Gas Production Facilities: A Multi-Year Empirical Study. *International Journal of Scientific Interdisciplinary Research*, *12*(10), 01–53. https://doi.org/10.63125/70msp878

[110]. Tahmedur, R., & Rifat, C. (2023). Critical Path Analysis and Resource-Leveling Optimization for Multi-Project Engineering Portfolios. *American Journal of Interdisciplinary Studies*, *13*(12), 01-47. https://doi.org/10.63125/8eqnea46

[111]. Tanoue, Y., Kawada, A., & Yamashita, S. (2017). Forecasting loss given default of bank loans with multi-stage model. *International Journal of Forecasting*, *33*(2), 513–522. https://doi.org/10.1016/j.ijforecast.2016.11.005

[112]. Tanoue, Y., & Yamashita, S. (2019). Loss given default estimation: A two-stage model with classification tree-based boosting and support vector logistic regression. *Journal of Risk*, *21*(4), 19–37. https://doi.org/10.21314/jor.2019.405

[113]. Tanoue, Y., Yamashita, S., & Nagahata, H. (2020). Comparison study of two-step LGD estimation model with probability machines. *Risk Management*, *22*(3), 155–177. https://doi.org/10.1057/s41283-020-00059-y

[114]. Tobback, E., Martens, D., Van Gestel, T., & Baesens, B. (2014). Forecasting loss given default models: Impact of account characteristics and the macroeconomic state. *Journal of the Operational Research Society*, *65*(3), 376–392. https://doi.org/10.1057/jors.2013.158

[115]. Wang, H., Forbes, C. S., Fenech, J.-P., & Vaz, J. (2020). The determinants of bank loan recovery rates in good times and bad: New evidence. *Journal of Economic Behavior & Organization*, *177*, 875–897. https://doi.org/10.1016/j.jebo.2020.06.001

[116]. Weber, P., Carl, K. V., & Hinz, O. (2024). Applications of explainable artificial intelligence in finance: A systematic review of finance, information systems, and computer science literature. *Management Review Quarterly*, *74*, 867–907. https://doi.org/10.1007/s11301-023-00320-0

[117]. Yao, X., Crook, J., & Andreeva, G. (2015). Support vector regression for loss given default modelling. *European Journal of Operational Research*, *240*(2), 528–538. https://doi.org/10.1016/j.ejor.2014.06.043

[118]. Ye, H., & Bellotti, A. (2019). Modelling recovery rates for non-performing loans. *Risks*, *7*(1), 19. https://doi.org/10.3390/risks7010019

[119]. Zhang, X., & Zhao, X. (2024). Using the Bayesian sampling method to estimate corporate loss given default distribution. *Journal of Empirical Finance*, *79*, 101540. https://doi.org/10.1016/j.jempfin.2024.101540