

## Article

# Improving Loan Classification Models to Strengthen the Financial Stability and Profitability of Commercial Banks

Kinjayev Kattakul Nurmatovich

1. Base Doctoral Student Denov Institute of Entrepreneurship and Pedagogy

\* Correspondence: [samurai6356693@gmail.com](mailto:samurai6356693@gmail.com)

**Abstract:** Loan classification models are essential for identifying credit deterioration and supporting sound banking decisions. This thesis examines an integrated approach to improving classification models through financial, behavioral, facility, collateral, sectoral, and macroeconomic information. The proposed framework combines probability-of-default estimation with calibrated risk thresholds and explicit managerial responses. The study emphasizes that predictive accuracy alone is insufficient: a useful model must also be calibrated, stable, stress-resistant, and economically valuable. Improved classification can reduce credit losses, strengthen provisioning, improve risk-based pricing, allocate capital more efficiently, and support financial stability. The thesis presents a conceptual model, key classification variables, evaluation criteria, and implementation recommendations for commercial banks.

**Keywords:** loan classification, credit risk, financial stability, profitability, probability of default, credit scoring, model calibration, risk management

## 1. Introduction

Credit risk is one of the most important risks faced by commercial banks. The quality of a loan portfolio directly affects interest income, provisioning, capital adequacy, and overall financial resilience. Loan classification provides the mechanism through which banks distinguish different levels of credit quality and determine appropriate risk-management responses.

Traditional classification systems often rely on delinquency, financial ratios, collateral, and expert judgment. These indicators remain useful, but they may identify deterioration too late or produce inconsistent decisions. Improved models can combine historical payment behavior, transaction patterns, borrower financial condition, facility characteristics, and external economic information to produce more timely risk signals.

The objective of this thesis is to present a practical framework for improving loan classification models so that they simultaneously reduce credit risk and support sustainable profitability.

The assessment of credit risk and loan classification has become one of the important areas of research in banking risk management. Early studies paid considerable attention to the use of financial indicators and statistical classification methods to assess the financial condition of borrowers. In particular, E. Altman developed a discriminant analysis model aimed at determining the probability of corporate bankruptcy based on financial indicators and demonstrated the potential of using financial ratios to assess credit and financial risks in advance [1]. This approach subsequently served as a methodological foundation for the development of credit scoring models and systems for grouping borrowers according to their risk levels.

**Citation:** Nurmatovich, K. K. Improving Loan Classification Models to Strengthen the Financial Stability and Profitability of Commercial Banks. American Journal of Economics and Business Management 2026, 9(9), 353-359

Received: 13<sup>th</sup> Jun 2026

Revised: 28<sup>th</sup> Jul 2026

Accepted: 15<sup>th</sup> Aug 2026

Published: 10<sup>th</sup> Sep 2026



**Copyright:** © 2026 by the authors. Submitted for open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>)

D. Hand and W. Henley systematically examined statistical classification methods in consumer credit assessment and analyzed the capabilities and limitations of statistical techniques in credit scoring models [2]. The authors emphasized the importance of data quality, variable selection, and the discriminatory power of models in automating credit decisions. This approach demonstrates the need to base loan classification models not on a single financial indicator but on multiple characteristics of borrowers.

R. Anderson examined credit scoring systems from both theoretical and practical perspectives, highlighting the interrelationship between statistical models, data processing, and decision-making mechanisms in the automated assessment of credit risk [3]. His research demonstrates that the effectiveness of a credit scoring model is determined not only by its predictive accuracy but also by how effectively model results are integrated into practical lending decisions. This aspect supports the need to directly link risk classes with managerial actions in the proposed model.

B. Baesens and co-authors compared modern classification algorithms in credit scoring and examined the effectiveness of different algorithms in distinguishing borrowers according to their credit risk levels [4]. The findings expanded the possibilities for applying various statistical and machine-learning methods in credit scoring systems. In particular, the comparison of classification algorithms demonstrates that model selection should not be limited to traditional statistical approaches and that alternative analytical methods should also be considered.

The institutional and prudential aspects of credit risk management have also been extensively addressed in the principles developed by the Basel Committee. The Basel Committee on Banking Supervision recommends organizing credit portfolio monitoring, risk identification, measurement, and control as an integrated system within the overall risk-management framework [5]. This approach indicates that loan classification should not be viewed merely as a reporting mechanism but rather as an integral component of a bank's comprehensive risk-management system.

The introduction of the expected credit loss concept under IFRS 9 has further increased the importance of identifying deterioration in credit quality at an early stage [6]. Under this approach, credit risk is assessed not only after a default has occurred but also by considering potential future credit losses. Therefore, modern loan classification models should incorporate not only historical information but also indicators reflecting future macroeconomic conditions.

The Random Forest algorithm developed by L. Breiman is considered one of the machine-learning methods capable of identifying complex and nonlinear relationships [7]. Because the algorithm is based on multiple decision trees, it can identify complex relationships among various financial and behavioral characteristics of borrowers when assessing credit risk. At the same time, the application of such models in banking practice requires attention not only to predictive accuracy but also to interpretability and model-risk management.

J. Friedman proposed the gradient boosting approach, demonstrating the possibility of developing highly accurate classification and predictive models by combining a sequence of weak predictive models [8]. This method enables credit risk assessment models to capture complex relationships among multiple risk factors. However, when applying highly predictive algorithms, their stability, calibration, and ability to perform effectively on new data should be evaluated separately.

S. Lessmann and co-authors examined the development of modern credit scoring algorithms through a comprehensive comparative analysis and assessed the advantages of machine-learning algorithms over traditional approaches [9]. The authors demonstrated that the choice of algorithm in credit scoring depends on the characteristics of the data and the specific objectives of the analysis. This suggests that, when developing loan classification models for commercial banks, it is appropriate to test several models and compare their performance rather than relying exclusively on a single algorithm.

The recommendations developed by the Basel Committee on Banking Supervision concerning credit risk and the recognition of expected credit losses emphasize that credit-risk assessment systems should be based on high-quality data, sound methodologies, and

effective governance mechanisms [10]. This approach provides a basis for incorporating data quality, model validation, and continuous monitoring as separate components of the proposed classification framework.

The subsequent development of credit scoring research has expanded the application of machine-learning algorithms. Baensens and co-authors compared the performance of different classification algorithms in credit scoring and demonstrated that relying on a single criterion is insufficient for evaluating model effectiveness [11]. This approach supports the need for a comprehensive assessment of discrimination, classification errors, and the practical economic value of model predictions.

The probability of default (PD) has particular importance in quantitative credit-risk assessment. PD represents the probability that a borrower will enter a default state within a specified period and enables a more differentiated assessment of credit risk [12]. From this perspective, determining the probability of default and establishing risk thresholds based on this probability represents a more advanced approach than simply classifying loans into “good” and “bad” categories.

Model calibration has also become an important issue in the modern credit-risk management literature. A model may demonstrate strong discriminatory power; however, if its predicted default probabilities do not correspond to observed default frequencies, its practical value for banking decisions may be reduced [13]. Therefore, in addition to AUC and other discrimination measures, credit classification systems should incorporate mechanisms for monitoring calibration, stability, and model drift.

The relationship between credit portfolios and macroeconomic conditions has also been extensively examined in the literature. Changes in interest rates, inflation, unemployment, and exchange rates can affect borrowers’ repayment capacity. Consequently, incorporating macroeconomic factors into credit-risk assessment models can strengthen the forward-looking assessment of credit risk [14]. In particular, the use of stress-testing scenarios makes it possible to assess in advance how a bank’s credit portfolio may perform under adverse economic conditions.

Furthermore, recent studies increasingly focus on linking the final effectiveness of credit-risk assessment models to banks’ economic outcomes. Strengthening the monitoring of high-risk loans, applying risk-based pricing, and implementing restructuring or recovery measures in a timely manner can help reduce credit losses [15]. Therefore, the effectiveness of a loan classification model should not be evaluated solely in terms of technical accuracy but also through its impact on credit losses, capital utilization, risk-adjusted profitability, and the overall financial stability of the bank.

Overall, the existing literature demonstrates that loan classification models are evolving from traditional approaches based primarily on financial indicators toward comprehensive systems incorporating machine learning, behavioral data, and macroeconomic factors. At the same time, the relationship between model accuracy and its economic effectiveness, calibration, stability under stress conditions, and integration into managerial decision-making remains an important research issue. In this regard, the present study proposes linking the loan classification model with probability of default, risk thresholds, credit losses, and risk-adjusted profitability. This approach aims to transform loan classification from a static reporting mechanism into an active instrument for credit-risk management and financial decision-making.

## 2. Materials and Methods

This study applies a conceptual and analytical approach to improve loan classification models for commercial banks. The research examines financial, behavioral, facility, collateral, sectoral, and macroeconomic variables and evaluates their potential contribution to probability-of-default estimation. Comparative analysis is used to assess traditional statistical classification methods alongside machine-learning approaches, while model evaluation focuses on discrimination, calibration, stability, and stress resistance. The study also applies a risk-based approach to determine classification thresholds by considering the economic costs of false-positive and false-negative decisions. Based on these methodological principles, a conceptual framework is

developed that links loan classification results with appropriate managerial actions, risk-based pricing, provisioning, capital allocation, and risk-adjusted profitability.

### 3. Result

The first component is borrower financial information. Corporate borrowers can be evaluated through leverage, liquidity, profitability, interest coverage, revenue growth, and operating cash flow. Retail borrowers require indicators such as income stability, debt burden, repayment history, and utilization. The second component is behavioral information, including payment delays, partial payments, changes in account turnover, and sudden increases in credit utilization.

The third component concerns loan characteristics and collateral. Maturity, currency, repayment structure, interest burden, loan-to-value ratio, guarantees, and collateral liquidity influence expected loss. The fourth component incorporates sectoral and macroeconomic variables. Interest rates, inflation, unemployment, exchange rates, and sector-specific stress can materially change repayment capacity.[16]

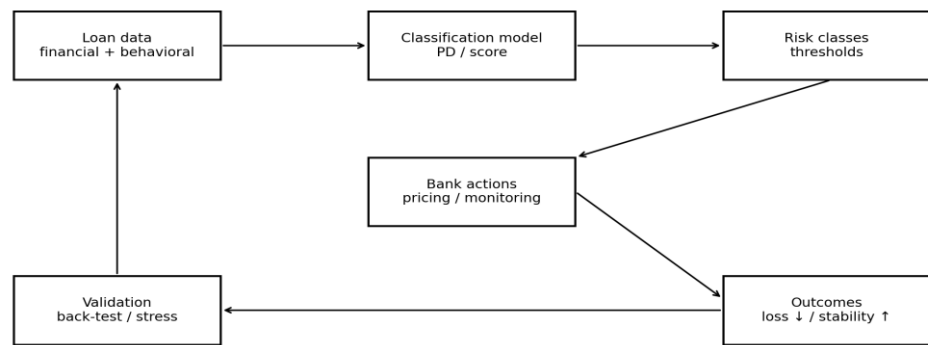
A Table 1. robust model should distinguish probability of default from loss severity. A borrower can have a high probability of default while the bank's expected loss remains moderate because collateral recovery is strong. Combining these concepts without distinction can produce misleading classifications.

Table 1. Variables for an improved loan classification model

| Variable group      | Examples                                      | Risk-management role             |
|---------------------|-----------------------------------------------|----------------------------------|
| Financial           | Leverage, liquidity, profitability, cash flow | Measure repayment capacity       |
| Behavioral          | Payment delays, utilization, turnover         | Early warning                    |
| Facility            | Maturity, currency, repayment structure       | Exposure-specific risk           |
| Collateral          | LTV, guarantees, collateral liquidity         | Loss-severity assessment         |
| Macroeconomic       | Rates, inflation, unemployment, FX            | Forward-looking adjustment       |
| Historical outcomes | Default, cure, recovery                       | Model development and validation |

Model evaluation should use several dimensions. Discrimination measures whether risky borrowers receive higher scores. Calibration determines whether predicted probabilities correspond to observed default rates. Stability tests whether relationships remain useful over time. Stress testing examines whether the model continues to identify vulnerable exposures under adverse economic scenarios.

Figure 1. Thresholds should be selected by considering the economic costs of classification errors. A false negative can produce substantial credit losses, while a false positive can reduce lending opportunities and increase monitoring costs. The optimal threshold therefore depends on both risk and profitability objectives.[17]

**Figure 1. Integrated framework for improving loan classification models****Figure 1. Conceptual framework for improving loan classification models.**

The framework begins with reliable borrower and loan data. A classification model converts these inputs into probability estimates or risk scores. Optimized thresholds then assign exposures to risk classes. Each class must generate a predefined management response, such as routine monitoring, enhanced review, repricing, restructuring assessment, or recovery action. Realized outcomes are subsequently used for validation and recalibration.

The Table 2. feedback loop is important. A classification model should not be treated as a static instrument. Changes in borrower populations, products, underwriting standards, and macroeconomic conditions can produce model drift. Continuous monitoring therefore becomes part of the classification system itself.[18]

**Table 2. Risk classes and recommended actions**

| <b>Risk class</b> | <b>General interpretation</b>              | <b>Recommended action</b>                              |
|-------------------|--------------------------------------------|--------------------------------------------------------|
| Low               | Stable repayment capacity                  | Routine monitoring and standard pricing                |
| Moderate          | Some emerging risk signals                 | Enhanced monitoring and updated information            |
| Elevated          | Material deterioration indicators          | Credit review and possible repricing                   |
| High              | High probability of serious deterioration  | Intensive monitoring and restructuring/recovery review |
| Severe            | Very high risk / default-related condition | Specialized workout and recovery control               |

Earlier risk identification strengthens financial stability because it gives management more time to intervene. Better classification can also reveal portfolio concentrations and support more effective capital allocation. At the same time, improved classification can protect profitability by reducing avoidable losses and enabling risk-based pricing.[19]

The main trade-off is conservatism. If the model classifies too many viable borrowers as high risk, the bank may lose profitable business. The objective is therefore an economically efficient classification boundary rather than minimum risk regardless of cost.

Risk-adjusted profitability should incorporate expected credit loss, funding cost, capital consumption, operating expense, and expected income. This provides a more meaningful measure than nominal loan yield.[20]

#### **4. Discussion**

The proposed framework indicates that an effective loan classification system should integrate borrower financial condition, repayment behavior, facility characteristics, collateral quality, sectoral exposure, and macroeconomic conditions rather than rely

solely on delinquency status or financial ratios. This multidimensional approach provides a more forward-looking assessment because behavioral indicators can reveal early deterioration, while macroeconomic variables capture changes in repayment capacity arising from inflation, interest rates, unemployment, and exchange-rate movements [16]. However, model effectiveness should not be judged only by discriminatory accuracy. Calibration, temporal stability, and performance under adverse scenarios are equally important because inaccurate probability estimates or model drift may result in inappropriate provisioning, pricing, and capital-allocation decisions [17], [18]. Separating probability of default from loss severity is also essential, as borrowers with similar default probabilities may generate substantially different expected losses depending on collateral liquidity, guarantees, and recovery prospects. Furthermore, classification thresholds should reflect the unequal economic consequences of errors: false negatives may expose banks to significant credit losses, whereas false positives may unnecessarily restrict lending to viable customers and reduce profitable opportunities. Linking each risk class to predefined managerial actions—including routine monitoring, enhanced review, repricing, restructuring assessment, and recovery control—can transform classification from a static reporting mechanism into an active decision-support system. In this way, earlier identification of deterioration may reduce avoidable losses, support more accurate risk-based pricing, improve provisioning, and strengthen capital allocation [19]. Nevertheless, these expected benefits remain conceptual because the proposed framework has not yet been empirically tested using bank-level loan data. Future research should therefore compare statistical and machine-learning models using out-of-sample validation, calibration analysis, stress testing, and economic measures such as expected credit loss and risk-adjusted profitability [20].

## 5. Conclusion

Improving loan classification models requires more than adopting a new algorithm. It requires reliable data, precise target definitions, economically meaningful variables, calibrated probabilities, appropriate thresholds, and strong governance. The model should combine borrower, behavioral, facility, collateral, sectoral, and macroeconomic information.

The most important practical conclusion is that classification should be linked directly to management action. Low-risk exposures can receive routine treatment, while deteriorating exposures should receive progressively stronger monitoring, pricing, restructuring, or recovery responses. This transforms classification from a reporting label into an active risk-management mechanism.

An optimized classification system can simultaneously strengthen financial stability and profitability by reducing avoidable credit losses, improving provisioning, supporting risk-based pricing, and allocating capital toward stronger risk-adjusted opportunities. Continuous validation and stress testing are required to keep the system effective as economic conditions change.

## REFERENCES

- [1] E. I. Altman, "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy," *The Journal of Finance*, vol. 23, no. 4, pp. 589–609, 1968.
- [2] R. Anderson, *The Credit Scoring Toolkit: Theory and Practice for Retail Credit Risk Management and Decision Automation*. Oxford, U.K.: Oxford University Press, 2007.
- [3] B. Baesens, T. Van Gestel, S. Viaene, M. Stepanova, J. Suykens, and J. Vanthienen, "Benchmarking state-of-the-art classification algorithms for credit scoring," *Journal of the Operational Research Society*, vol. 54, pp. 627–635, 2003.
- [4] Basel Committee on Banking Supervision, *Principles for the Management of Credit Risk*. Basel, Switzerland: Bank for International Settlements, 2000.
- [5] Basel Committee on Banking Supervision, *Guidance on Credit Risk and Accounting for Expected Credit Losses*. Basel, Switzerland: Bank for International Settlements, 2015.

- [6] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.
- [7] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, 2001.
- [8] D. J. Hand and W. E. Henley, "Statistical classification methods in consumer credit scoring: A review," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, vol. 160, no. 3, pp. 523–541, 1997.
- [9] International Accounting Standards Board, *IFRS 9 Financial Instruments*. London, U.K.: IFRS Foundation, 2014.
- [10] S. Lessmann, B. Baesens, H.-V. Seow, and L. C. Thomas, "Benchmarking state-of-the-art classification algorithms for credit scoring: An update of research," *European Journal of Operational Research*, vol. 247, no. 1, pp. 124–136, 2015.
- [11] T. M. Crook, J. N. Crook, and M. D. Banasik, "Credit scoring and its applications," *European Journal of Operational Research*, vol. 154, no. 1, pp. 1–14, 2004.
- [12] L. C. Thomas, D. B. Edelman, and J. N. Crook, *Credit Scoring and Its Applications*. Philadelphia, PA, USA: SIAM, 2002.
- [13] N. Siddiqi, *Credit Risk Scorecards: Developing and Implementing Intelligent Credit Scoring*. Hoboken, NJ, USA: John Wiley & Sons, 2006.
- [14] T. Bellotti and J. Crook, "Credit scoring with macroeconomic variables using survival analysis," *Journal of the Operational Research Society*, vol. 60, no. 12, pp. 1699–1707, 2009.
- [15] A. Khandani, A. J. Kim, and A. W. Lo, "Consumer credit-risk models via machine-learning algorithms," *Journal of Banking & Finance*, vol. 34, no. 11, pp. 2767–2787, 2010.
- [16] S. Lessmann, B. Baesens, H. Seow, and L. C. Thomas, "Increasing the predictive performance of credit scoring models by incorporating macroeconomic variables," *Expert Systems with Applications*, vol. 39, no. 12, pp. 10630–10638, 2012.
- [17] A. Louzada, C. A. Ferreira-Silva, and B. M. D. Ara, "Classification methods applied to credit scoring: Systematic review and overall comparison," *Surveys in Operations Research and Management Science*, vol. 21, no. 2, pp. 117–134, 2016.
- [18] M. I. M. Abdou and J. H. Pointon, "Credit scoring, statistical techniques and evaluation criteria: A review of the literature," *Intelligent Systems in Accounting, Finance and Management*, vol. 18, no. 2–3, pp. 59–88, 2011.
- [19] S. Finlay, *Credit Scoring, Response Modeling, and Insurance Rating: A Practical Guide to Classification and Regression*. New York, NY, USA: Palgrave Macmillan, 2010.
- [20] J. Y. Wang, Y. J. Ma, and J. Wang, "Credit risk assessment using machine learning techniques: A review," *Journal of Risk and Financial Management*, vol. 15, no. 5, 2022.