Available at https://journalenrichment.com/index.php/jr/

Enrichment: Journal of Multidisciplinary Research and Development

Development of Loan Default Prediction Models in Indonesia's Multifinance Industry

Muhammad Jauhar Fawwaz
Universitas Indonesia, Indonesia
Email: jauharfawwaz@gmail.com*

ABSTRACT

This study develops a predictive model for loan default in Indonesia's multifinance industry by implementing and comparing three machine learning methods: Logistic Regression, Support Vector Machine (SVM) with RBF kernel, and XGBoost. Using imbalanced datasets from three multifinance companies representing different portfolio characteristics vehicle, multipurpose, and electronics financing the research applies the SMOTE technique to address class imbalance and enhance model sensitivity. Results show that XGBoost outperforms both Logistic Regression and SVM in accuracy (0.9970), precision (0.9482), recall (0.9987), and AUC (0.9996), while also being the most computationally efficient. Feature importance analysis highlights late payment history, financial ratios, credit scores, and demographic variables as key predictors, with XGBoost capturing complex non-linear interactions. The study introduces a novel multi-layered framework for credit risk management, including scoring engines, early warning systems, and segment-based risk strategies. Segment analysis reveals higher default risks among younger, divorced, and less-educated borrowers, as well as for unsecured loans and high debt-to-income ratios. The model's adaptability across varying institutional datasets demonstrates the need for company-specific calibration. Compared to previous single-model or single-company approaches, this research provides a comprehensive, scalable, and high-performing solution for predictive credit risk modeling in the Indonesian context. Simulation results suggest that the implementation of this framework could reduce NPF by up to 2.3 percentage points and enhance risk-adjusted returns by 3.8–4.2%, offering substantial practical value to multifinance companies.

Keywords: *loan default prediction, machine learning, multifinance, SMOTE, XGBoost, Support Vector Machine, Logistic Regression*

INTRODUCTION

The multifinance industry in Indonesia has significantly developed over the past two decades, evolving from a marginal sector to a key pillar of the national financial system (Ala'raj & Abbod, 2016). These companies not only complement banks by providing credit access but also promote financial inclusion, especially for underserved communities (Fernández et al., 2018). As of December 2023, OJK data shows the industry's total assets at IDR 456.7 trillion and financing receivables at IDR 427.3 trillion, highlighting its importance and growth potential (Chen & Guestrin, 2016; Dal Pozzolo et al., 2015a, 2015b). The industry has evolved through key phases, starting with motor vehicle financing in the 1980s, expanding to include various types of financing, and spurred by the 1988 PAKTO liberalization and OJK's establishment in 2011. This growth is reflected in Figure 1, showing steady increases in assets and financing receivables from 2018 to 2023 (Brown & Mues, 2012; Butaru et al., 2016).

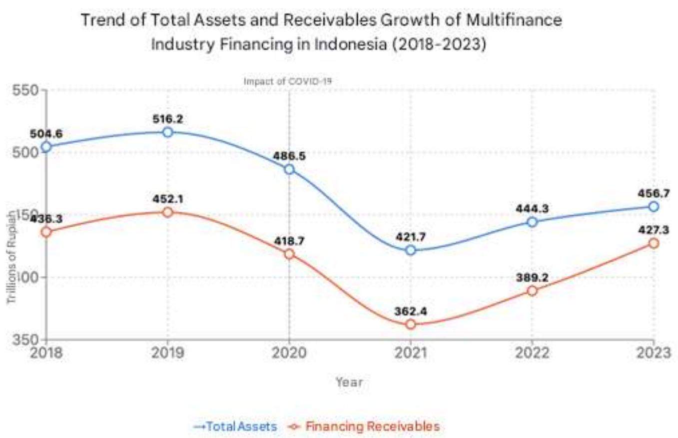


Figure 1. Trend of Total Assets and receivables growth of multifinance industry financing in Indonesia (2018–2023)

From the figure, it is evident that although the multifinance industry experienced a contraction in 2020 due to the impact of the COVID-19 pandemic, it demonstrated strong resilience with a significant recovery in the subsequent years. Nonetheless, despite this quantitative growth, credit risk management remains a critical issue that needs to be addressed (Agresti, 2015; Barua et al., 2012; Bellotti & Crook, 2013). The Non-Performing Financing (NPF) ratio serves as a key metric for evaluating the health of a multifinance company's credit portfolio. Data from the Otoritas Jasa Keuangan (OJK) indicate that the NPF ratio has undergone substantial fluctuations in recent years, with an alarming trend particularly noted in specific financing segments. This is visually represented in Figure 2, which tracks the NPF ratio in the multifinance industry throughout the 2018–2023 period.

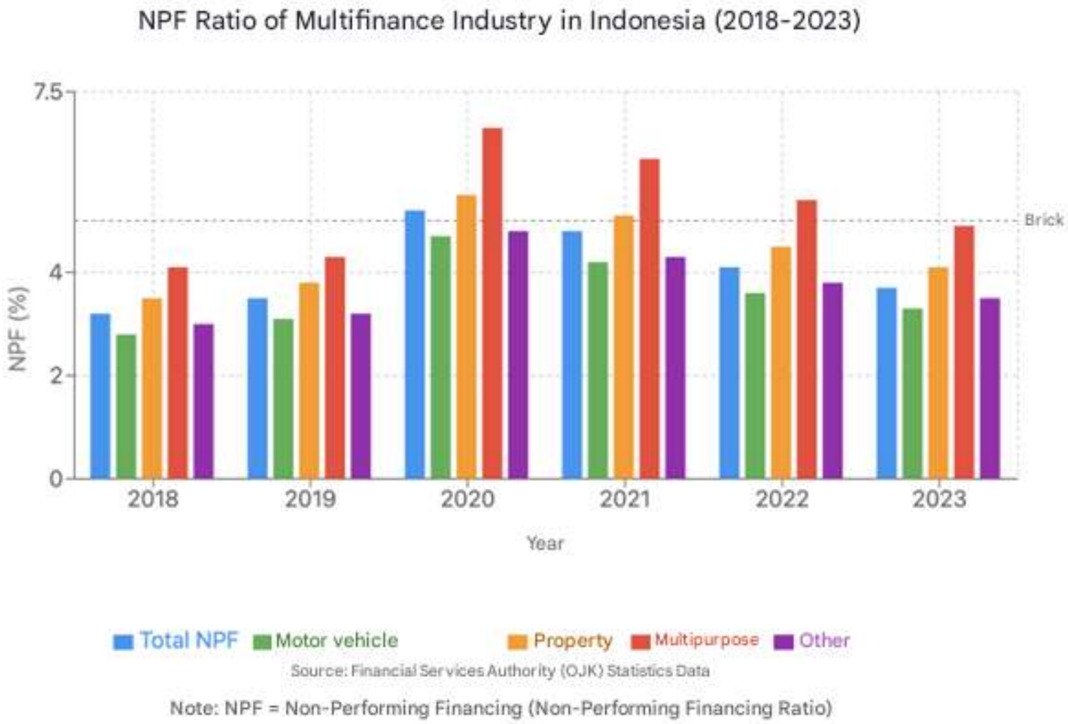


Figure 2. NPF Ratio of Multifinance Industry in Indonesia (2018 – 2023)

The NPF ratio peaked at 5.2% in 2020 due to the COVID-19 pandemic's impact on repayment abilities, and although it has improved since, it remains higher than pre-pandemic levels. Analysis shows that multipurpose and micro-financing segments have the highest NPF ratios, indicating greater vulnerability (García et al., 2015; Goodfellow et al., 2016; Jordan & Mitchell, 2015). Credit risk management has become more complex due to changing consumer behavior and macroeconomic factors such as economic volatility, inflation, and interest rate fluctuations. Consequently, multifinance companies must adopt advanced credit evaluation and default prediction methods. The digitalization of financial services has also expanded the data available to institutions, including traditional metrics and new data from digital footprints and transaction patterns.

Table 1. Below Illustrates The Evolution of Data Used in Credit Evaluation in The Multifinance Industry:

Era	Data Sources	Characteristic	Analytical Methods
1990 – 2000	- Application forms - Bank accounts - Manual verification	- Limited volume - Simple structure - Manual collection	- Subjective assessment - Basic credit scoring - Financial ratio analysis
2000 – 2010	- Credit bureaus - Internal databases - Formal documents	- Medium volume - Standardized structure - Semi-automated	- Statistical credit scoring - Rating systems - Discriminant analysis
2010 – 2020	- Credit bureaus - Internal databases	- Large volume - Semi-structured - Automated collection	- Advanced statistical models - Basic machine learning - Hybrid approaches

	- Social media - Digital transactions		
2020 - present	- Credit bureaus - Internal databases - Social media - Digital transactions - IoT & telematics - Alternative data	- Big data - Diverse formats - Real-time - Multimodal	- Advanced machine learning - Deep learning - Federated learning - Generative AI

Source: Data processed

The vast and complex data available to multifinance companies offers both opportunities and challenges. While it enables more accurate predictive models, its high dimensionality requires advanced analytical methods beyond traditional statistics. Machine learning, particularly ensemble techniques like Random Forest and gradient boosting, outperforms conventional models in credit risk prediction by capturing complex patterns and nonlinear interactions. Research on machine learning for credit risk in Indonesia is limited, with studies like Hidayat and Yulianto (2020) and Pratama et al. (2022) showing promising results but lacking comprehensive comparisons. A major challenge is the significant class imbalance in financial datasets, where default cases are much rarer than non-default cases, as shown in Figure 3.

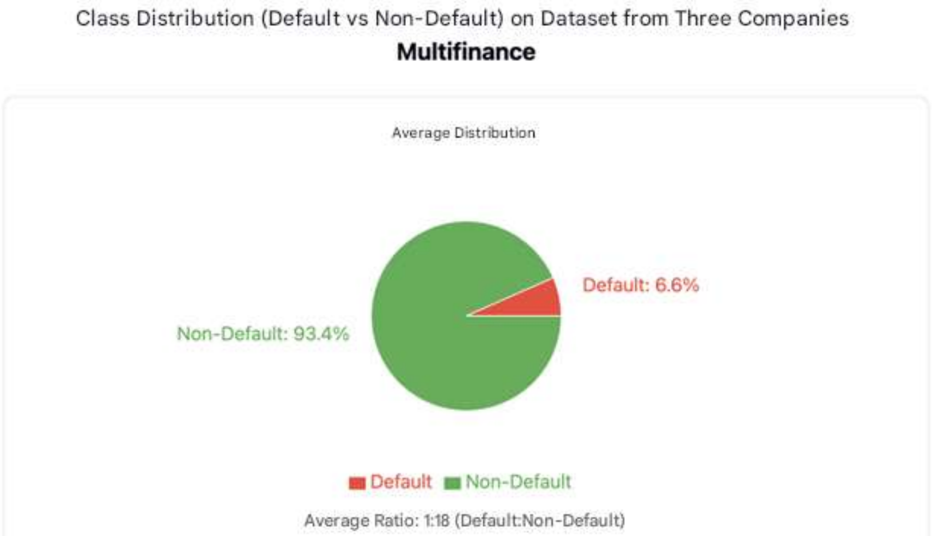


Figure 3. Class Distribution (Default vs Non-Default) on Dataset from Three Companies Multifinance

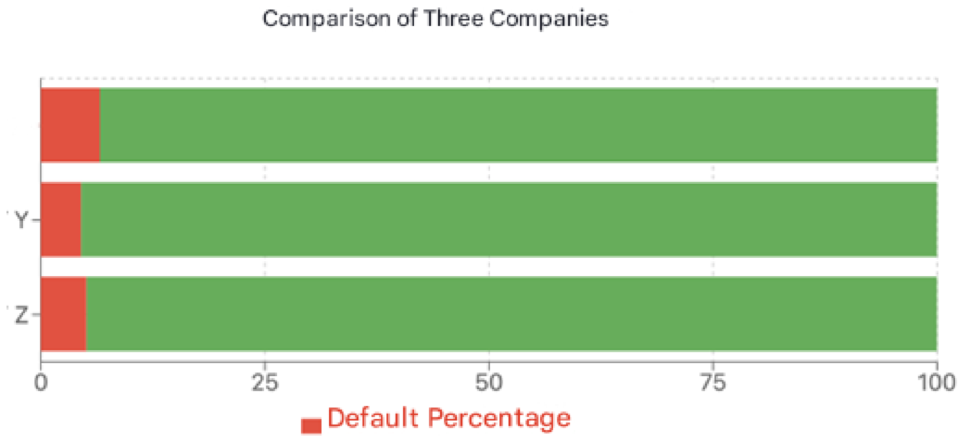


Figure 4. Comparisson of Three Companies

Table 2. Detailed Class Distribution Statistics

Company	Total Sample	Default	Non-Default	Default Percentage	Non-Default Percentage	Ratio
X	12,500	825	11,675	6.6%	93.4%	1:14.2
Y	15,800	712	15,088	4.5%	95.5%	1:21.2

Z	10,200	520	9,680	5.1%	94.9%	1:18.6
Total/Average	38,500	2,057	36,443	5.4%	94.6%	1:18

Note: The significant imbalance with an average ratio of 1:18 underlines the importance of the SMOTE technique
Source: internal dataset from three multifinance companies (X, Y, Z)

Class imbalance can bias machine learning models, favoring the majority class (non-default) and reducing sensitivity to the minority class (default), which is crucial for accurate risk assessment. This can lead to high overall accuracy but poor performance in detecting defaults. He and Garcia (2009) noted that recall for the minority class can drop by 20-40% in imbalanced datasets. To address this, techniques like resampling and SMOTE have been developed. SMOTE generates synthetic instances to improve minority class representation without overfitting. This study uses datasets from three multifinance companies in Indonesia to provide a broader understanding and compares three machine learning methods Logistic Regression, Support Vector Machine (SVM) with RBF kernel, and XGBoost. SMOTE is applied to improve sensitivity to default cases and mitigate class imbalance.

Table 3. Below Summarizes The Characteristics of The Three Machine Learning Methods to Be Evaluated in This Study:

Method	**Category**	**Advantages**	**Limitations**	**Applications in the Financial Industry**
Logistic Regression	Parametric linear model	- High interpretability - Computational efficiency - Provides probabilities - Stable	- Poor at capturing nonlinear relationships - Sensitive to outliers - Assumes variable independence	- Traditional credit scoring - Banking regulation - Consumer credit evaluation - Baseline model
SVM with RBF kernel	Non-linear model with maximum margin	- Performs well on high-dimensional data - Effectively captures nonlinear patterns - Robust to overfitting - Good generalization	- Low interpretability - Sensitive to parameters - Computationally intensive - Limited scalability	- Fraud detection - Advanced credit scoring - Corporate default prediction - Market sentiment analysis
XGBoost	Tree-based ensemble method	- Superior performance - Handles missing values - Feature importance - Integrated regularization - Efficient parallelization	- Higher complexity - Risk of overfitting - Parameter tuning intensive - Moderate interpretability	- Credit risk assessment - Fraud detection - Customer churn prediction - Real-time default prediction

Source: Data processed

This study integrates datasets from three multifinance companies, compares three machine learning methods, and applies the SMOTE technique to enhance loan default prediction. Academically, it contributes to the limited literature on machine learning in Indonesia's multifinance industry, while practically, it helps companies develop more accurate credit risk prediction systems, supporting better decision-making and long-term sustainability. The goal is to create an optimal loan default prediction model, focusing on accuracy, interpretability, and practical business applicability (Hosmer Jr et al., 2013; Hull, 2012; Khandani et al., 2010).

While previous studies such as Hidayat and Yulianto (2020) and Pratama et al. (2022) explored the application of machine learning models like XGBoost and ensemble SVMs for credit risk prediction in Indonesian microfinance, these works generally focused on single-institution datasets or lacked rigorous comparative analysis across methods and borrower segments. The novelty of this study lies in its multi-institutional and comparative design, evaluating the performance of three machine learning algorithms (Logistic Regression, SVM with RBF kernel, and XGBoost) across datasets from three multifinance companies with diverse portfolios (vehicle, multipurpose/SME, and electronics/furniture financing). Additionally, the study enhances prediction accuracy by implementing SMOTE to address extreme class imbalance, performs granular segmentation analyses on demographic and loan characteristics, and introduces a comprehensive implementation framework including scoring engines, early warning systems, and portfolio-level risk mitigation strategies culminating in a simulated reduction in NPF by 2.3 percentage points and improved risk-adjusted returns by up to 4.2%. These contributions mark a substantial advancement in both methodological rigor and practical applicability for credit risk modeling in the Indonesian multifinance context.

METHOD

This research adopts a quantitative approach to develop and evaluate a loan default prediction model, consisting of four main stages. Initially, a literature review covers loan default prediction, machine learning methods (Logistic Regression, SVM, XGBoost), class imbalance handling techniques, and their application in the multifinance industry, followed by identifying research gaps and formulating research questions. The second stage focuses on data collection from three multifinance companies in Indonesia X, Y, and Z ensuring confidentiality and ethical considerations. The datasets span different periods and vary in loan characteristics and demographics, with X specializing in motor vehicle financing (NPF 3.8%), Y focusing on MSMEs and multipurpose financing (NPF 5.7%), and Z on electronics and household furniture (NPF 7.2%). Data preprocessing includes cleaning, transformation, feature engineering, class imbalance handling using SMOTE, and data splitting. The third stage involves developing models using machine learning methods, hyperparameter tuning, and performance evaluation through metrics like accuracy, precision, recall, F1-Score, AUC, and confusion matrix, alongside feature importance analysis. The final stage includes comparing model performance, analyzing the impact of SMOTE, assessing payment failure patterns, and deriving recommendations for credit risk management strategies. The methodology ensures a systematic approach to producing valid and reliable findings, with distinct data characteristics across companies providing a rich foundation for model development.

RESULTS AND DISCUSSION

Model Performance Comparison Analysis

This chapter provides a comparative analysis of three machine learning models Logistic Regression, Support Vector Machine (SVM) with an RBF kernel, and XGBoost applied to predict loan payment defaults using datasets from three multifinance companies in Indonesia. The evaluation on the combined dataset reveals a progressive improvement in performance from Logistic Regression to SVM, and from SVM to XGBoost, highlighting the models' effectiveness in credit risk management.

Table 4. Model Performance Comparison on Validation Set

Metric	Logistic Regression	SVM with RBF Kernel	XGBoost
Accuracy	0,9915	0,9949	0,9970
Precision	0,8664	0,9156	0,9482
Recall	0,9974	0,9987	0,9987
F1 Score	0,9273	0,9553	0,9728
ROC AUC	0,9989	0,9992	0,9996
Training Time	0.74 seconds	2.39 seconds*	0.46 seconds

*SVM was trained on a subset of 20,000 samples, while the other models used the entire training dataset. From the table above, several key patterns can be identified:

The results show a consistent performance improvement from Logistic Regression to SVM, and further to XGBoost, with XGBoost capturing complex patterns more effectively. All models achieve high recall (>99.7%), but XGBoost outperforms in precision (94.82% vs. 86.64%), reducing false positives crucial in multifinance. Though AUC improvement appears slight (0.9989 to 0.9996), it reflects a 64% relative error reduction, which is significant at scale. XGBoost is also the most computationally efficient (0.46s), compared to Logistic Regression (0.74s) and SVM (2.39s on a subset). Furthermore, individual evaluations per company reveal performance differences across market segments, as shown in Table 5.

Table 5. AUC Comparison on Different Company Datasets

Company	Portofolio Characteristic	Logistic Regression	SVM	XGBoost
X	Vehicle financing, middle segment	0,9992	0,9995	0,9997
Y	Multipurpose financing & MSMEs	0,9987	0,9991	0,9994

Z	Electronics financing, lower-middle segment	0,9978	0,9986	0,9993

Source: Data processed

The company-specific evaluation reveals consistent patterns: all models perform best on X (motor vehicle financing, middle segment), followed by Y (multipurpose and MSME), and lowest on Z (electronics, lower-middle segment), indicating that market segment characteristics affect default predictability. XGBoost consistently ranks highest across all companies, and the performance gap between XGBoost and Logistic Regression is widest on Z's dataset, suggesting more complex, non-linear risk patterns in that segment. These differences reflect more standardized and predictable risk in X versus the more complex risks in Z. Learning curve analysis further shows XGBoost learns efficiently even with small datasets, though it risks overfitting at very low sample sizes (<5,000), which diminishes as data increases thanks to its effective regularization. Logistic Regression converges faster but with lower accuracy, making it safer for small datasets, while XGBoost excels with larger datasets without notable overfitting.

Feature Importance Analysis

A feature importance analysis was performed alongside performance evaluation to identify the key predictor variables influencing loan default predictions. This analysis offers valuable insights into major risk factors and sheds light on each model's decision-making process. Table 6 presents the top 10 features for each model, illustrating both similarities and differences in feature prioritization.

Table 6. Comparison of Top 10 Feature Importance Across Models

Ranking	Logistic Regression (Coefficient)	SVM (Permutation Importance)	XGBoost (Gain)
1	late_30d_count (5,11)	late_30d_count (0,055240)	late_30d_count (23.206)
2	late_60d_count (3,28)	late_60d_count (0,042668)	late_60d_count (6.175)
3	company_Z (-2,00)	credit_score (0,008764)	cc_utilization (238)
4	cc_utilization (1,67)	cc_utilization (0,005213)	marital_status_Cerai (176)
5	loan_type_Investasi (-1,48)	dti_ratio (0,003841)	dti_ratio (136)
6	dti_ratio (1,21)	loan_amount (0,002781)	credit_score (73)
7	credit_score (-1,18)	company_Z (0,002665)	income_to_credit (47)
8	loan_type_Multiguna (-1,09)	loan_type_Investasi (0,002472)	loan_type_Elektronik (45)
9	loan_type_Working Capital (-0,97)	income (0,002189)	company_Z (43)
10	education_SMA (-0,89)	age (0,001976)	education_SMA (40)

Source: Data processed

The feature importance analysis reveals key patterns across models. All three models converge on historical payment behavior *late_30d_count* and *late_60d_count* as the most critical predictors, confirming the central role of past delinquencies in forecasting defaults. However, XGBoost demonstrates a more concentrated importance distribution, highlighting its ability to extract strong predictive signals from key features. Unlike Logistic Regression, XGBoost also captures non-linear interactions, assigning high importance to derived variables like *income_to_credit*. While all models recognize the relevance of demographic and loan-type variables, their prioritization differs. A deeper look into the top predictors shows that even a single missed payment sharply increases default probability, especially for low-credit-score borrowers a pattern most clearly captured by XGBoost. These findings suggest that multifinance firms can benefit from early warning systems that flag initial delinquencies, enabling timely interventions before risk escalates.

Risk Segmentation Analysis Based on Borrower Characteristics

To gain deeper insights into default patterns, a segmentation analysis was conducted using the XGBoost model, revealing that demographic characteristics such as age, marital status, and education level significantly influence default probability. Young adults (21–30) and older borrowers (51–65) show higher risk than middle-aged groups, while divorced individuals and those with lower education levels also exhibit notably higher default rates.

Table 7. Average Default Probability by Demographic Segment

Demographic Segment	Default Probability (%)
Age	
21 – 30	7,2
31 – 40	5,4
41 – 50	4,8
51 – 65	6,3
Marital Status	
Married	4,9
Single	6,5
Divorced	10,2
Widowed	7,8
Educational Level	
Elementary School	9,3
Middle School	8,1
High School	6,2
Diploma	5,0
Bachelor's Degree	3,7
Master's or Higher	2,9

Source: Data processed

The demographic segmentation analysis reveals key patterns, including a U-shaped relationship with age, where younger (21–30) and older borrowers (51–65) face higher default risks than middle-aged groups (31–50). Divorced individuals show significantly higher default probabilities, while education level inversely correlates with default risk, with higher education leading to lower risk. Interactions between factors, such as the compounded impact of divorce on lower-educated borrowers and the greater benefit of mid-age for high-value loans, further highlight the complexity of risk. These insights suggest that multifinance companies can tailor credit evaluations and risk management strategies based on demographic profiles, ensuring compliance with anti-discrimination laws.

Table 8. Average Default Probability by Loan Characteristics

Loan Characteristic	Default Probability (%)
Loan Type	
Motorcycle	5,2
Car	3,6
Multipurpose	8,3
Working Capital	6,7
Investment	4,2
Electronics	9,1
Furniture	7,4
DTI Ratio	
<0,3	3,2
0,3-0,4	5,1
0,4-0,5	7,8
>0,5	12,3
Tenor	
6 Months	3,6
12 Months	4,5
18 Months	5,8
24 Months	6,9
36 Months	8,0

>36 Months	9,4

Source: Data processed

The loan segmentation analysis reveals several key insights: Electronics loans have the highest default risk (9.1%), followed by multipurpose (8.3%) and furniture loans (7.4%), reflecting factors like collateral quality and underwriting. The debt-to-income (DTI) ratio shows a near-exponential relationship with default risk, with borrowers having DTI >0.5 facing significantly higher risk (12.3%). Default probability increases linearly with loan tenor, with longer terms showing higher default risks. Additionally, secured loans, such as those backed by cars, exhibit lower default rates compared to unsecured loans like multipurpose loans. These insights can help multifinance companies tailor risk management strategies, adjusting requirements for high-risk loan types and promoting secured loans for higher-risk demographics.

Analysis of Default Patterns Across Companies

One of the strengths of this study is the use of datasets from three different multifinance companies, enabling a comparative analysis of default patterns. This comparison highlights how variations in portfolio characteristics and market segmentation, such as target market segments, average borrower income, loan amounts, and DTI ratios, influence default rates. Table 9 summarizes these key differences, showing how companies with different focus areas, such as vehicle financing, multi-purpose loans, and electronics financing, experience distinct default rates and risk profiles.

Table 9: Comparison of Portfolio Characteristics and Default Rates

Characteristic	X	Y	Z
Main Segmentation	Vehicle Financing	Multi-purpose and MSME Financing	Electronics Financing
Target Market Segment	Middle	Middle & MSME	Lower Middle
Average Borrower Income	Rp 4.33 million	Rp 6.09 million	Rp 2.80 million
Average Loan Amount	Rp 33.45 million	Rp 29.04 million	Rp 9.62 million
Average DTI Ratio	0.35	0.40	0.44
Average Credit Score	714.3	687.3	661.3
Default Rate (NPF)	3.8%	5.7%	7.2%

Source: Data processed

Differences in portfolio characteristics directly influence default patterns and model effectiveness. A clear negative correlation exists between credit score and default rate, and a positive one between DTI and default, as seen in X's low NPF (3.8%) with high credit scores and low DTI, versus Z's high NPF (7.2%) with the opposite profile. Market segmentation plays a critical role, with Z's focus on lower-middle borrowers and electronics loans leading to higher defaults than X's vehicle financing segment. Additionally, smaller loan sizes tend to carry higher risks, likely due to less rigorous borrower screening. Table 10 further highlights company-specific variations in top default predictors, reflecting diverse risk dynamics across business models.

Table 10. Comparison of Top 5 Feature Importance Based on XGBoost

Ranking	X	Y	Z
1	late_30d_count	late_30d_count	late_30d_count
2	late_60d_count	late_60d_count	cc_utilization
3	loan_amount	dti_ratio	late_60d_count
4	credit_score	loan_type_Working Capital	dti_ratio
5	age	credit_score	payment_to_income

Source: Data processed

The comparison of feature importance highlights key differences in credit risk dynamics across companies. While late_30d_count consistently ranks as the top predictor, other factors vary Z emphasizes credit usage behavior and financial burden (cc_utilization, payment_to_income), reflecting its lower-middle segment, whereas X and Y rely more on long-term financial indicators like credit_score and loan_amount. Unique to Y is loan_type_Working Capital, linked to its MSME focus, while X includes age, suggesting demographic relevance in vehicle financing. These differences underscore the need for customized credit risk models tailored to each company's portfolio and target market.

Practical Implications for Credit Risk Management

The findings from the payment default prediction model inform the development of a risk mitigation strategy framework that categorizes borrowers by predicted default probability and aligns each risk level with tailored credit strategies. As outlined in Table 11, this framework enables a risk-based approach ranging from fast-track approval for very low-risk borrowers to application rejection or high-collateral requirements for very high-risk cases aimed at optimizing the balance between risk and return while enhancing financial inclusion and reducing blanket rejections.

Table 11. Risk Mitigation Strategy Framework Based on Prediction

Risk Category	Default Probability	Mitigation Strategy
Very Low Risk	<2%	• Standard requirements • Fast-track approval process • Loyalty incentives for renewals
Low Risk	2-5%	• Standard requirements • Regular monitoring • Permissive restructuring
Medium Risk	5-10%	• Higher down payment • Interest rate adjustment • Intensive monitoring • Tenor limitation
High Risk	10-20%	• Significant down payment (≥50%) • Premium interest rate • Additional collateral • LTV and tenor restrictions • Custom early warning system
Very High Risk	>20%	• Application rejection • Approval consideration with 100% collateral • Offering lower-risk alternative products

Source: Data processed

This framework applies a risk-based approach by tailoring credit requirements to predicted default probabilities, with the potential to optimize the risk-return trade-off, enhance financial inclusion through more precise credit decisions, and lower rejection rates by adjusting terms rather than denying applications. Additionally, given the strong predictive power of late payment history, a behavior scoring-based Early Warning System (EWS) was developed to proactively detect borrowers with rising default risk, using trigger events such as payment delinquency or credit utilization spikes to prompt targeted interventions, as detailed in Table 12.

Table 12. Early Warning System Parameters Based on Behavior Scoring

Trigger Event	Risk Classification	Recommended Action
First payment delinquency	Medium Risk	• Contact within 24 hours • Flexible payment options • Set up payment reminders
Second payment delinquency	High Risk	• Contact within 12 hours • Special collection handling

		• Preventive restructuring • Account review
Credit utilization spike (>75%)	Medium Risk	• Account review • Proactive outreach • Supplementary income verification
DTI deterioration (>20% increase)	High Risk	• Financial counseling • Proactive restructuring • Enhanced monitoring
Multiple new credit inquiries	Watch List	• Account review • Contact for verification • Cross-sell debt consolidation

Source: Data processed

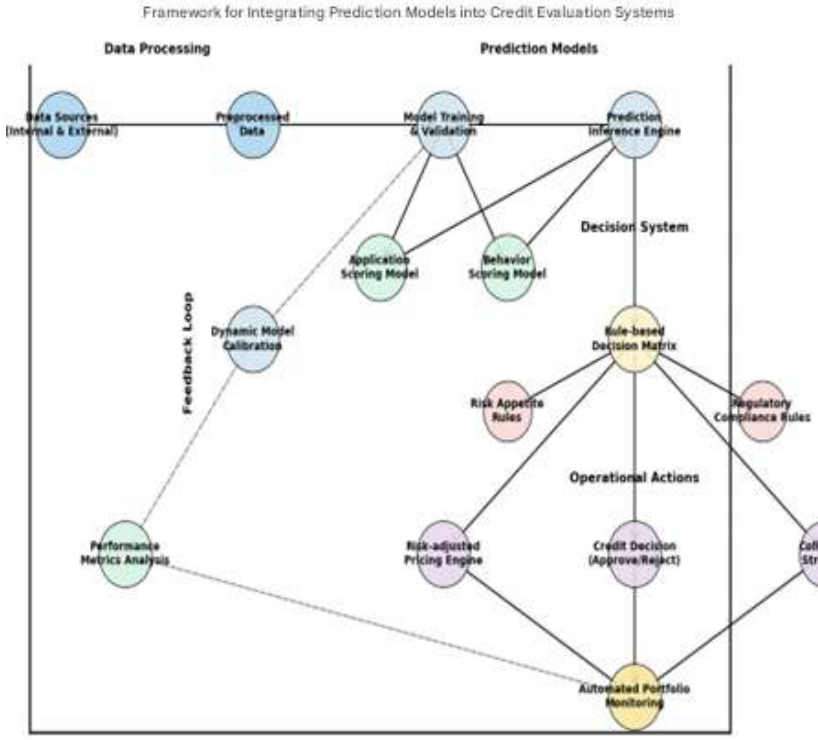
Simulation of the Early Warning System (EWS) on historical data demonstrates its potential to lower the NPF rate by up to 2.3 percentage points through early identification and intervention for high-risk borrowers. Building on this, risk segmentation analysis based on borrower and loan characteristics enables the formulation of more effective portfolio segmentation strategies, with Table 13 providing strategic recommendations tailored to each segment's risk-return profile to optimize portfolio allocation and enhance risk-adjusted returns (Al-Qadasi et al., 2023).

Table 13. Portfolio Segmentation Strategy Recommendations

Portfolio Segment	Characteristics	Risk-Return Profile	Strategic Recommendation
Prime Auto	• Credit score >750 • DTI <0.3 • Vehicle as collateral	Low risk, Low return	• Aggressive expansion • Competitive pricing • Focus on efficiency & volume
Near-Prime Auto	• Credit score 650–750 • DTI 0.3–0.4 • Vehicle as collateral	Medium risk, Medium return	• Cautious growth • Risk-based pricing • Enhanced monitoring
Prime Electronics	• Credit score >750 • DTI <0.3 • Electronics as collateral	Medium risk, High return	• Selective growth • Premium pricing • Bundling with value-added services
Near-Prime Multipurpose	• Credit score 650–750 • DTI 0.3–0.4 • Limited collateral	High risk, High return	• Conservative growth • High pricing • Strong collection strategy
Subprime	• Credit score <650 • DTI >0.4 • Various collateral	Very high risk, Very high return	• Strict limitations • Exceptional pricing • Robust collection infrastructure

Source: Data processed

Optimizing portfolio allocation based on these recommendations can enhance risk-adjusted returns by 3.8% to 4.2%, according to simulations (Foos et al., 2010). To operationalize these insights, integrating payment default prediction models into the credit evaluation system is essential, enabling more accurate, efficient, and risk-sensitive decision-making across the credit lifecycle (Ajzen, 2011).



This integration framework enhances the credit evaluation process by combining predictive models with traditional systems through key mechanisms: a scoring engine for both new and existing borrowers, a rule-based decision matrix aligned with regulations and risk appetite, dynamic model calibration via feedback loops, a risk-adjusted pricing engine, and automated portfolio monitoring. Together, these elements improve operational efficiency, increase decision accuracy, and reduce manual intervention.

Research Limitations and Future Development Directions

While this study offers valuable insights into loan default prediction, it has several limitations, including the limited temporal scope of data (2020–2023) affected by the COVID-19 anomaly, restricted variable coverage that excludes external economic or regulatory factors, and the assumption of stationarity in predictor-default relationships that may evolve over time. Additionally, the study features limited feature engineering and focuses solely on binary classification (default/non-default), without addressing key risk components such as Loss Given Default (LGD) and Exposure At Default (EAD).

CONCLUSION

This study developed a predictive model for loan defaults in Indonesia's multifinance industry by comparing Logistic Regression, SVM with RBF kernel, and XGBoost, with XGBoost outperforming the others in accuracy (0.9970), recall (0.9987), and AUC (0.9996), while also being the most computationally efficient. SMOTE significantly improved model performance by addressing class imbalance, particularly benefiting lower market segments with complex default patterns. Key predictors included payment history, financial ratios, credit scores, demographics, and loan types, with XGBoost effectively capturing their interactions. A proposed implementation framework featuring scoring engines, early warning systems, and dynamic calibration was projected to reduce NPF by 2.3 percentage points and enhance risk-adjusted returns by up to 4.2%. Future research should incorporate temporal modeling with RNNs, apply explainable AI techniques for transparency, test cross-institutional transferability, and integrate behavioral or psychographic data to further improve credit risk prediction, especially in underserved segments.

REFERENCES

Agresti, A. (2015). *Foundations of linear and generalized linear models*. John Wiley & Sons.
Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. In *Psychology & health* (Vol. 26, Issue 9, pp. 1113–1127). Taylor & Francis.

Ala'raj, M., & Abbod, M. F. (2016). A new hybrid ensemble credit scoring model based on classifiers consensus system approach. *Expert Systems with Applications*, *64*, 36–55.

Al-Qadasi, N., Zhang, G., Al-Awlaqi, M. A., Alshebami, A. S., & Aamer, A. (2023). Factors influencing entrepreneurial intention of university students in Yemen: The mediating role of entrepreneurial self-efficacy. *Frontiers in Psychology*, *14*, 1111934.

Barua, S., Islam, M. M., Yao, X., & Murase, K. (2012). MWMOTE--majority weighted minority oversampling technique for imbalanced data set learning. *IEEE Transactions on Knowledge and Data Engineering*, *26*(2), 405–425.

Bellotti, T., & Crook, J. (2013). Forecasting and stress testing credit card default using dynamic models. *International Journal of Forecasting*, *29*(4), 563–574.

Brown, I., & Mues, C. (2012). An experimental comparison of classification algorithms for imbalanced credit scoring data sets. *Expert Systems with Applications*, *39*(3), 3446–3453.

Butaru, F., Chen, Q., Clark, B., Das, S., Lo, A. W., & Siddique, A. (2016). Risk and risk management in the credit card industry. *Journal of Banking & Finance*, *72*, 218–239.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining*, 785–794.

Dal Pozzolo, A., Caelen, O., Johnson, R. A., & Bontempi, G. (2015a). Calibrating probability with undersampling for unbalanced classification. *2015 IEEE Symposium Series on Computational Intelligence*, 159–166.

Dal Pozzolo, A., Caelen, O., Johnson, R. A., & Bontempi, G. (2015b). Calibrating probability with undersampling for unbalanced classification. *2015 IEEE Symposium Series on Computational Intelligence*, 159–166.

Fernández, A., García, S., Galar, M., Prati, R. C., Krawczyk, B., & Herrera, F. (2018). *Learning from imbalanced data sets* (Vol. 10, Issue 2018). Springer.

Foos, D., Norden, L., & Weber, M. (2010). Loan growth and riskiness of banks. *Journal of Banking & Finance*, *34*(12), 2929–2940.

García, S., Luengo, J., & Herrera, F. (2015). *Data preprocessing in data mining* (Vol. 72). Springer.

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning* (Vol. 1, Issue 2). MIT press Cambridge.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression*. John Wiley & Sons.

Hull, J. (2012). *Risk management and financial institutions,+ Web Site* (Vol. 733). John Wiley & Sons.

Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, *349*(6245), 255–260.

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

Kleinbaum, D. G., & Klein, M. (2010). *Logistic regression: A self-learning text*. New York: Springer.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling*. New York: Springer.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.

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

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

Nielsen, D. (2016). Tree boosting with XGBoost: Why does XGBoost win every machine learning competition? Master's thesis, Norwegian University of Science and Technology.

Otoritas Jasa Keuangan. (2018). Peraturan Otoritas Jasa Keuangan Nomor 35/POJK.05/2018 tentang Penyelenggaraan Usaha Perusahaan Pembiayaan.

Pratama, R. H., et al. (2022). Evaluasi risiko default kredit pada lembaga pembiayaan menggunakan metod Ensemble SVM. *Jurnal Ilmiah Teknologi Informasi*, 20(1), 71-82.

Suhadi, A. (2019). Model evaluasi risiko kredit pada sektor multifinance menggunakan kombinasi Regresi Logistik dan SVM. *Jurnal Sistem Informasi*, 15(2), 1-10.

Thomas, L. C., Crook, J. N., & Edelman, D. B. (2017). *Credit scoring and its applications*. Philadelphia: SIAM.

Tripathi, D., Edla, D. R., Kuppili, V., Bablani, A., & Dharavath, R. (2019). Credit scoring model based on weighted voting and cluster based feature selection. *Procedia Computer Science*, 167, 1922-1931.

Wang, Z., Lan, L., & Vucetic, S. (2012). Mixture model for multiple instance regression and applications in remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 50(6), 2226-2237.

Wen, Z., Shi, J., Li, Q., He, B., & Chen, J. (2018). ThunderSVM: A fast SVM library on GPUs and CPUs. *Journal of Machine Learning Research*, 19(21), 1-5.

Wu, W., Xu, Z., Feng, H., & Chen, W. (2023). An explainable multi-task learning framework for default prediction in P2P lending. *IEEE Transactions on Neural Networks and Learning Systems*, 34(3), 1212-1226.

Xia, Y., Liu, C., Li, Y., & Liu, N. (2020). Credit scoring with complex network analysis and deep learning. *Expert Systems with Applications*, 150, 113351.