

Predicting Loan Default Among Borrowers Using Bootstrapped Digital Lending Platforms in Metro Manila, Philippines: A Survey-Based Gradient Boosting Approach

Neil T. Awit^{1*}, Allan A. Calderon², Emmanuel J. Dotong³

^{1,2} College of Business Administration, Lyceum of the Philippines University, Manila, Philippines

³ Ramon V. Del Rosario College of Business De La Salle University, Manila, Philippines

Received : September 16,
2026

Revised: October 28,
2026

Accepted : October 28, 2026

Online : October 30, 2026

Abstract

This study developed a predictive model of loan default among 512 borrowers using bootstrapped digital lending platforms in Metro Manila, Philippines. Data were collected using a validated survey instrument via quota-controlled stratified random sampling. Quotas were allocated based on the bootstrapped digital lending platforms respondents used. CatBoost was used to predict loan repayment outcomes from self-reported survey data. Model evaluation employed a 70/30 training-test split and 5-fold cross-validation on the full dataset. Feature importance analysis and SHapley Additive exPlanations (SHAP) were used to interpret the model's predictions. Loan characteristics accounted for 90.81% of the model's reported feature importance, with loan amount alone accounting for 82.85%. SHAP analysis further indicated that higher loan amounts were generally associated with higher predicted default risk. Behavioral traits and financial capacity indicators accounted for 2.91% and 6.28% of the reported feature importance, respectively. These findings were leveraged in the development of credit risk assessment practices among bootstrapped digital lending platforms.

Keywords: *Financial Technology (Fintech), Gradient Boosting, Loan Default, Predictive Modeling, CatBoost, SHAP, Credit Risk Assessment*

INTRODUCTION

With the continuous innovation-induced market disruptions, the business landscape in the Philippines has been evolving over the last decade. Significant advancements and increased adoption of analytics technologies have been observed. Specifically in terms of digitalization, various traditional brick-and-mortar businesses are now shifting to the use of e-commerce platforms, allowing them to further expand the potential of their operations, profitability, and the overall health of their business. At the same time, this shift is allowing other related industries to thrive.

One of those is Financial Technology (Fintech), a domain within the financial services industry that leverages technology to provide various financial solutions. Thus, its rapid growth trajectory is projected to continue in the next few years, if not in the next few decades, as digital adoption in the country is still growing. As reported by [Statista \(n.d\)](#), there were already 289 registered Fintech platforms operating in the country as of 2022, with an average growth of 9%. Furthermore, digital payments made by individuals have grown significantly by 47,642, translating to 63.1% versus non-digital payments, as presented in the Bangko Sentral ng Pilipinas (BSP)'s report on the status of digital payments in the country.

Copyright Holder:

© Neil. Allan, & Emmanuel. (2026)

Corresponding author's email: neil.awit@lpu.edu.ph

This Article is Licensed Under:



Given the vast scope of solutions these Fintech platforms can offer, as well as what the previous literature suggests, this paper focuses on the personal loan business of these Fintech platforms. Specifically, this paper focuses on bootstrapped digital lending platforms, given that this niche market is a game changer, although it is hounded by challenges, including loan defaults. Bootstrapped digital lending platforms refer to small-scale fintech companies that operate with limited external funding (Jagtiani & Lemieux, 2019). Hence, they are severely impacted by loan defaults. In fact, according to Tonik Bank (2024), non-performing loans (NPLs) or bad loans hit P427 billion as of April 2023, of which P1.4 billion, or 7.43%, come from digital banks, which include bootstrapped digital lending platforms.

Although existing credit-risk assessment mechanisms use borrower- and loan-related information to screen applicants, the reviewed literature indicates that predictive modeling has been applied predominantly in conventional banking and established lending environments. Previous studies have used logistic regression, decision trees, random forests, and gradient boosting to examine credit scoring, loan approval, and repayment prediction (Shinde et al., 2022; Dhruv et al., 2023). However, its application, especially in the Philippine context, remains underexplored. In addition, examining behavioral traits, financial capacity indicators, and loan characteristics within a common predictive framework may provide a comparative assessment of their contributions to model predictions. Hence, to address these gaps, this paper examined gradient boosting algorithms for predicting loan repayment outcomes among borrowers using bootstrapped digital lending platforms in Metro Manila, Philippines. This is pivotal in providing context-specific evidence that may inform the development of credit-risk assessment practices.

Objectives of the Study

- To predict loan repayment outcomes among borrowers on bootstrapped digital lending platforms using the top-performing machine learning (ML) algorithm.
- To rank and examine the predictive contributions of financial capacity indicators, loan characteristics, and behavioral traits to loan repayment outcomes using feature importance analysis and SHapley Additive exPlanations (SHAP).
- To provide data-driven credit risk assessment and mechanisms for bootstrapped digital lending platforms.

LITERATURE REVIEW

Loan Default Prediction and Traditional Credit-Risk Modeling

This paper focuses on the application of predictive analytics, particularly Machine Learning (ML), not only to the loan screening process but also to the in-depth analysis of the factors associated with borrowers' repayment and default behavior. Loan default prediction is fundamentally a credit-risk assessment problem, in which borrower characteristics, financial conditions, and loan-related information are examined to estimate the likelihood of repayment failure (Purnawan & Danu Prasetyo, 2021).

This approach is closely aligned with the study by Shinde et al. (2022), which applied logistic regression to the loan screening process in the context of banks in India. Sheik et al. (2020) used logistic regression to predict the loan approval process in banking, based on dummy data. Anand et al. (2022) also used logistic regression to predict customers' tendency to default on their loans. Manglani and Bokhare (2021) applied the same predictive model in the specific context of retail banking to determine whether loan delinquency is likely. In the lending industry, Dutta (2021) applied machine learning to predict the likelihood of customers repaying their loans. Hence, these

studies establish the application of predictive modeling in credit assessment, specifically through logistic regression and other supervised classification approaches, as these models generally rely on borrower and loan information to distinguish among borrowers who may default. This suggests that traditional variables typically used for loan screening may be suitable for studies where binary classification algorithms such as logistic regression are applied (Jian Ai et al., 2021). However, the increasing interest in borrower-level behavioral and financial data provides an opportunity to extend loan default prediction beyond standard approaches (Faraji, 2022). As Adamu (2018) postulates, digital lending environments are particularly relevant for borrowers' financial and usage behaviors, providing additional information for credit-risk assessment (Kumar et al., 2021).

Machine Learning Approaches to Loan Default Prediction

It is very apparent that the application of predictive analytics beyond logistic regression in the banking sector is increasing due to its perceived value and potential to uncover non-traditional variables (Patiu & Eleazar, 2024). For example, Dhruv et al. (2023) examined income, demographics, and creditworthiness to predict bank loan repayments. Meanwhile, Kumar et al. (2021) used deep learning models to predict the banking system's loan scoring mechanism, providing a multidimensional and comprehensive analysis of customer loan behavior. Furthermore, Madaan et al. (2021) used black-box algorithms such as decision trees and random forest classifiers to predict customer loan defaults. Similarly, Dansana et al. (2024) used the same algorithm, such as a random forest classifier, to predict customer loans. Meanwhile, Kokate and Chetty (2021) applied various complex algorithms, such as decision trees, gradient boosting, and random forest classifiers, to predict the loan default tendencies of customers in the commercial bank sector. Meanwhile, Prasanth et al. (2023) focused on the use of a random forest classifier to predict customers' loan eligibility and therefore approval.

These findings suggest a growing interest in applying ML to behavioral finance studies, which can capture nonlinear relationships and generate more complex patterns in credit-related data (Kokate & Chetty, 2021). Specifically, the growing application of tree-based ensemble methods, such as Random Forest and Gradient Boosting, provides alternatives to conventional linear or parametric approaches for credit-risk prediction. Hence, more complex predictive models, particularly gradient boosting models, advance studies related to customer loan behavior. Gradient boosting algorithms are particularly relevant because they build an ensemble of sequentially improved decision trees, allowing the model to capture complex relationships between predictors and loan outcomes. Kokate and Chetty (2021) incorporated gradient boosting, specifically decision trees and random forest classifiers, to predict loan default tendencies, demonstrating the applicability of boosting-based approaches to credit-risk prediction while highlighting the need to examine their application in more specific lending contexts. Hence, the application of complex ML models introduces an interpretability concern, as models such as gradient boosting can be difficult to interpret compared with conventional statistical approaches (Sheik et al., 2020). Also, interpretability becomes important when the objective is not only to predict default but also to reveal the contributing factors to the prediction (Harjanto & Nurim, 2023).

Financial Capacity and Loan Characteristics as Predictors of Loan Default

Loan default prediction can also be understood through the borrower's characteristics and the loan itself. First, financial capacity indicators describe the borrower's ability to meet financial obligations, while loan characteristics are attributes of the borrowing arrangement that may be associated with repayment outcomes (Lee & Kim, 2020). For example, Dhruv et al. (2023) examined customers' income, demographics, and creditworthiness in predicting bank loan repayments, revealing that these variables supported integrating borrower-level financial and demographic

information into predictive models to comprehensively assess repayment performance. [Dutta \(2021\)](#) also applied machine learning to predict the likelihood of customers' loan repayments, demonstrating the application of predictive analytics to borrower repayment behavior.

Furthermore, the use of loan behavior as a factor in classification models is evident in studies such as [Shinde et al. \(2022\)](#), [Sheikh et al. \(2020\)](#), and [Manglani and Bokhare \(2021\)](#). These studies examined loan screening, approval, or repayment outcomes, modeling them with borrower- and loan-related information. The general findings served as a basis for examining financial capacity indicators and loan characteristics as separate predictors rather than treating all predictors as a single group. Hence, they distinguish the applicability of these domains for studies that aim to examine financial capacity, loan characteristics, and behavioral traits and their varying contributions to predicting loan default ([Harjanto & Nurim, 2023](#)).

Behavioral Dimensions in Loan Default and Digital Lending

Beyond standard financial and loan-related variables, borrower behavior also provides another source of information for predicting loan default ([Adiani, 2021](#)). This is because these variables can capture patterns that are not necessarily reflected in conventional measures of creditworthiness or financial capacity ([Aattouf et al., 2021](#)). Hence, this paper focuses on the application of predictive analytics, particularly Machine Learning (ML), not only to the loan screening process but also to the in-depth analysis of borrowers' behavioral patterns, which are referred to as "observations" in this paper. Furthermore, the increasing application of digital financial services generates behavioral information that is particularly relevant, as digital platforms can produce data related to users' financial activities and engagement, which can be complementary to traditional borrower and loan characteristics in credit-risk assessment ([Arowolo et al., 2022](#)).

[Aattouf et al. \(2021\)](#) combined supervised and unsupervised learning algorithms. Specifically, supervised learning algorithms were used to predict the e-signing of the loans, while the unsupervised learning algorithms assisted in segmenting the customers based on their banking or loan behaviors. [Hegde et al. \(2022\)](#) also used clustering, association, and classification algorithms to predict the loan approval process of the banks. [Arowolo et al. \(2022\)](#) utilized multiple algorithms to predict bank loans, using Agglomerative Hierarchical Clustering (AHC) together with Decision Trees and Random Forest Classifiers to analyze the loan behaviors of the customers.

Gradient Boosting in Bootstrapped Digital Lending

From the previous analyses, it was revealed that mostly supervised algorithms were used; the question is whether black-box or white-box algorithms will be applied. There are also studies conducted in the past that utilize unsupervised learning algorithms. One is the study of [Hegde et al. \(2022\)](#), where they used clustering, association, and classification algorithms to predict the loan approval process of the banks. [Arowolo et al. \(2022\)](#) also utilized multiple algorithms to predict bank loans using Agglomerative Hierarchical Clustering (AHC), together with Decision Trees and Random Forest Classifiers in analyzing the loan behaviors of the customers. Furthermore, [Aattouf et al. \(2021\)](#) combined supervised and unsupervised learning algorithms, specifically, supervised learning algorithms were used to predict the e-signing of the loans while the unsupervised learning algorithms assisted in the segmentation of the customers based on their banking or loan behaviors.

Theoretical Framework

Furthermore, this study adopted information asymmetry as the theoretical basis. This model is crucial for understanding why borrower and loan characteristics may be informative in credit-risk assessment. In lending, borrowers typically hold more information about their financial

circumstances than lenders can directly observe, making the assessment of creditworthiness especially important for lending decisions. In this regard, there are three groups of predictors: behavioral traits, financial capacity indicators, and loan characteristics. These predictors emerged from the synthesis of the literature and the circumstantial relevance of the adopted theory. Together, these variables provide different forms of information that may contribute to predicting loan repayment outcomes.

Conceptual Framework

Figure 1 presents the conceptual framework, illustrating the interrelationships among predictors, including behavioral traits, financial capacity, and loan characteristics, as well as the outcome variable, loan repayment. Each construct is represented by its corresponding measured variables. This further implies that this study examines the predictive contributions of these constructs and does not imply causal relationships.

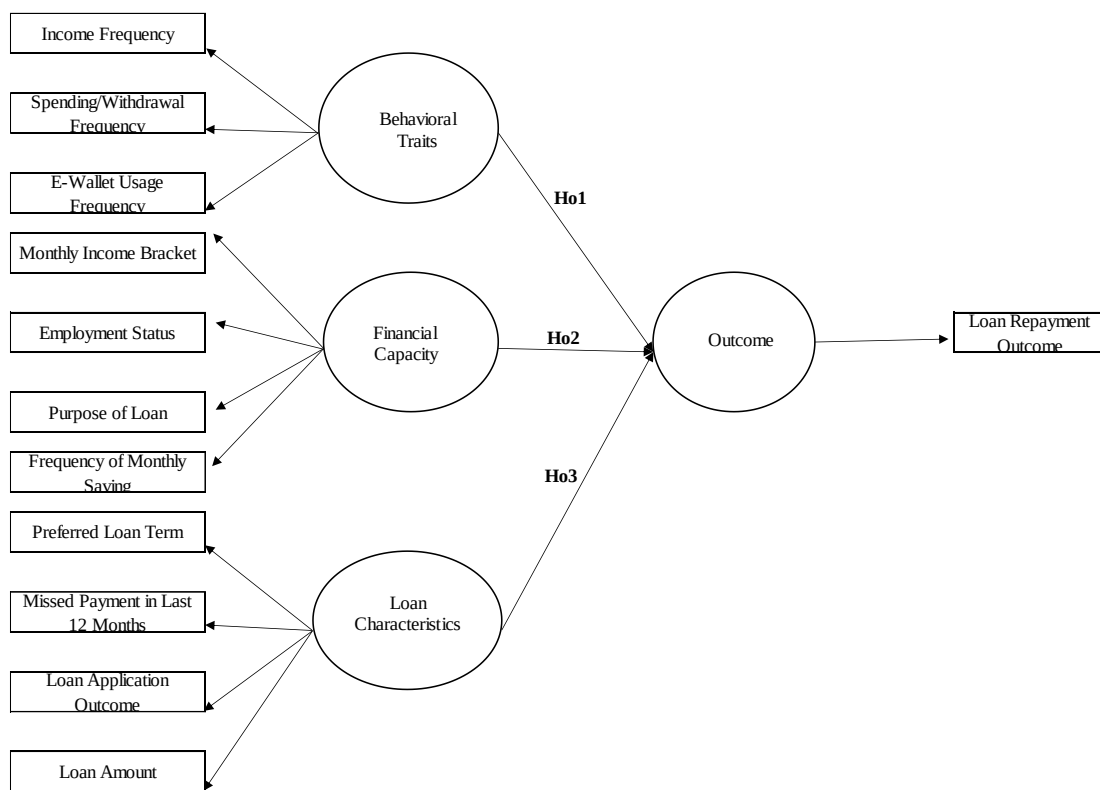


Figure 1. Conceptual Framework

RESEARCH METHOD

Research Design, Sampling, and Data Collection

This study employed a quantitative approach because data collection was survey-based. ML was used to predict loan repayment outcomes among borrowers using bootstrapped digital lending platforms in Metro Manila, Philippines. Accordingly, the target population comprised borrowers who had used digital lending applications offering personal loans. Quota sampling via an online questionnaire recruited a total of 512 respondents. Preliminary survey responses and available information identified eight commonly used bootstrapped digital lending platforms, which served as the basis for setting the quotas. A quota of 64 respondents per platform was established, yielding a total sample of 512 respondents. Equal allocation was intended to obtain comparable numbers of

observations across the eight platforms. However, this allocation was a quota-based sampling decision and should not be interpreted as proportional representation of platform usage in the target population. Respondents were recruited through social media groups, personal networks, and online communities used by borrowers. Responses were monitored by the platform reported by each respondent. Once the quota of 64 eligible respondents for a platform was reached, additional responses associated with that platform were excluded from the final sample. The identities of the platforms were anonymized to preserve confidentiality and data privacy, as required by the platforms themselves. Hence, the findings are presented at the aggregate level.

Furthermore, eligible respondents, excluding students, must have experience with the relevant digital lending services and currently reside in Metro Manila, which provides a geographically bounded setting for examining borrowers' financial circumstances and repayment outcomes. In addition, these respondents must have reported one completed loan to capture their repayment process and experience. This also allowed the repayment outcome to be classified as either fully repaid or not fully repaid, serving as the label for the predictive model used in this paper. Prior to the administration of the survey instrument, it underwent expert review and pilot testing. Google Forms were used to administer the survey. Participation was voluntary, and informed consent was obtained from respondents. Responses were collected anonymously, and the information provided was treated confidentially. Appropriate measures were also taken to safeguard collected data and ensure their proper disposal when necessary.

Variables and Data Preparation

The final dataset contained 11 predictor variables and one outcome variable. The predictors were organized into three conceptual categories: behavioral traits, financial capacity indicators, and loan characteristics. The outcome variable represented loan repayment status, coded as 0 for borrowers who did not fully repay their loans and 1 for borrowers who fully repaid their loans. The dataset was screened for duplicate and invalid responses, and none were identified in the final dataset. Furthermore, categorical variables were encoded as required for model implementation. Lastly, the Synthetic Minority Over-sampling Technique (SMOTE) was used to address class imbalance in the outcome variable.

Predictive Modeling and Model Evaluation

The main pipeline of this study is to compare different gradient boosting algorithms, such as CatBoost, XGBoost, and LightGBM, with logistic regression as a baseline model. These were implemented in Python using the Scikit-learn and CatBoost libraries. A 70/30 training-test split was used for model development and evaluation, and then hyperparameter tuning using GridSearchCV and five-fold cross-validation was conducted. The best-performing model was selected based on these evaluation results. Furthermore, feature importance analysis was conducted to examine the relative importance and contribution of predictors to the selected model's predictions. This was supplemented by SHapley Additive exPlanations (SHAP) to interpret model predictions and how individual features contributed to predicted outcomes. These combined techniques were used to reveal and describe predictive patterns.

FINDINGS AND DISCUSSION

As part of the pre-processing technique, the Synthetic Minority Oversampling Technique (SMOTE) was used to prevent data imbalance and, consequently, bias in the results, given that Class 0 (not repaid) comprised 39.9% while Class 1 (repaid) made up 60.1% of the total. Next, Table 1 shows the different ML models tested in this study. In terms of Accuracy, Logistic Regression performs best, while XGBoost is the least effective. For Precision, Logistic Regression is also the best

performer, while LightGBM is the least effective. For Recall, LightGBM shows an impressive result, while CatBoost performs the least. For F1-Score, both XGBoost and CatBoost perform the least, while LightGBM is the best performer. Lastly, AUC-ROC shows that CatBoost performs the best, while XGBoost performs the least. Given these results, if interpretability and stability are the emphasis, then Logistic Regression should be the model. However, the study emphasizes predictive performance for detecting defaulters, so, based on AUC-ROC, CatBoost is the selected model, as it is highly effective with the categorical-heavy data used in this study. This aligns with previous studies, which suggest that CatBoost is excellent with categorical-heavy data (Kwofie et al., 2015; Chen et al., 2021). Despite CatBoost's AUC-ROC of 0.81 being only marginally higher than LightGBM's 0.77, CatBoost was chosen for its superior handling of categorical-heavy data, which, at the same time, is capable of minimizing preprocessing complexity. Hence, the model is well-suited to the dataset used in this study and is practically feasible for a real-world application such as loan screening.

Table 1. Model Performance Comparison

Model	Accuracy	Precision	Recall	F1-Score	AUC-ROC
XGBoost	0.68	0.74	0.71	0.73	0.74
LightGBM	0.70	0.68	0.86	0.76	0.77
Logistic Regression	0.73	0.77	0.74	0.75	0.77
CatBoost	0.70	0.79	0.68	0.73	0.81

Note: Precision, recall, and F1-score are binary classification metrics calculated with Class 1 (fully repaid) as the positive class. These are class-specific metrics, not macro-averaged or weighted-averaged scores. AUC-ROC is calculated with respect to the repayment class, while accuracy represents the proportion of correctly classified observations.

Table 2 shows the variables and their individual effects on CatBoost's predictive power. First, behavioral traits collectively exhibit weak importance, contributing only 2.91%, indicating that the subvariables of behavioral traits are not significant drivers in the model's predictions. Next, the financial capacity indicators contribute slightly more than the behavioral traits, at a collective importance of 6.28%. This suggests that the subvariables of financial capacity have only minor influence on the model. Finally, the customers' loan characteristics emerged as the most influential predictors, collectively at 90.81%. Specifically, the loan amount is the most dominant variable within the loan characteristics, accounting for 82.85% of the model's prediction.

Given the results, borrowers aiming for larger loan amounts showed a higher tendency toward loan default. This may be due to borrowers overestimating their future repayment ability, particularly if their income is unstable (Bastani et al., 2019). Also, when customers apply for larger loan amounts, it may be driven by temporary negative shocks, such as financial distress (Chen et al., 2021). These findings align with previous studies (Kwofie et al., 2015; Dastile et al., 2020), reinforcing the importance of loan amount as a predictor of customers' inherent risk of default.

Table 2. Feature Importance

Constructs	Variables	Importance	Total Importance	Interpretation
Loan Characteristics	Loan Amount	82.853	90.806	Most dominant predictor group
	Missed Payment in Last 12 Months	4.878		
	Preferred Loan Term	1.279		
	Loan Application Outcome	1.796		

Financial Capacity	Monthly Income Bracket	1.692	6.285	Minor contribution
	Employment Status	1.830		
	Purpose of Loan	0.308		
	Frequency of Monthly Saving	2.454		
Behavioral Traits	Income Frequency	2.322	2.910	Very Minor contribution
	Spending/Withdrawal Frequency	0.588		
	E-Wallet Usage	0.000		
	Frequency			

In addition to Feature Importance, which explains the individual effects of variables on CatBoost's predictive power, SHAP (SHapley Additive exPlanations) was also used to add a layer of interpretability to CatBoost, as shown in Figure 2. While Feature Importance analysis identifies the relative contribution of each variable to the model, SHAP provides additional information on the direction and magnitude of each variable's contribution to predictions, complementing rather than replacing the Feature Importance analysis.

First, the blue dots on the left side of the Loan Amount plot suggest that borrowers with smaller loan amounts are less likely to default. This indicates that higher loan amounts are generally associated with higher predicted default risk, while smaller loan amounts are associated with lower predicted default risk. On the other hand, the various colors on the right side suggest that borrowers with larger loan amounts have a higher risk of default. Hence, the directional pattern aligns with the dominant feature-importance result for loan amount, as shown in Table 2. Next, the concentration of dots in various colors suggests that their influence on the model is much weaker than that of the Loan Amount. This is consistent with the low feature importance score shown in Table 2. Hence, the comparatively smaller SHAP contributions of the other variables indicate lesser influence on the model's individual predictions relative to loan amount.

This explains that the loan amount is indeed a crucial driver, driven by exposure risks (Serrano-Cinca et al., 2015). In other words, a higher loan amount entails a higher risk of default because it typically requires longer repayment terms, and longer duration entails greater uncertainty and therefore higher risk (Zhang et al., 2022). The results are consistent with prior studies (Zhang et al., 2022; Serrano-Cinca et al., 2015; Ndayishimiye & Kim, 2021). It is important to note, though, that these SHAP results should be interpreted as model-based predictive associations rather than causal effects. The analysis indicates how the loan amount contributed to CatBoost's predictions within the observed dataset but does not establish that increasing the loan amount itself causes borrowers to default.

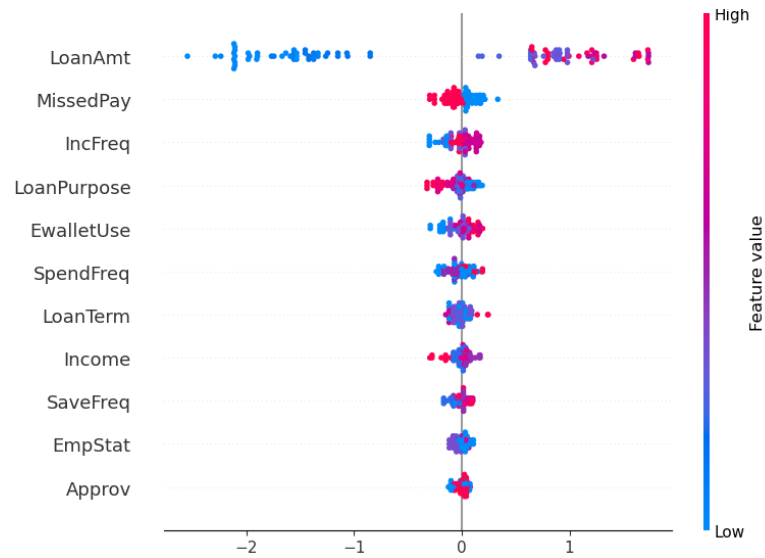


Figure 2. SHAP Beeswarm Plot

Also, Figure 3 shows the SHAP Waterfall Plot, which analyzes how each variable contributed to the prediction for a specific borrower. The red bar for LoanAmt, corresponding to ₱20,000 and a contribution of +0.88, indicates that the loan amount contributed positively to the model's predicted outcome for the class being explained. This suggests that a specific loan amount can substantially influence the individual prediction compared with other variables included in the model. Furthermore, the waterfall plot can be interpreted as an observation-level explanation rather than as a general relationship, in contrast to the SHAP beeswarm plot, which summarizes the contribution of predictors across observations. Thus, the waterfall analysis provides additional interpretability to the CatBoost model by showing the individual contributions of the predictors, while the global Feature Importance and SHAP beeswarm analyses provide evidence regarding the broader contribution of the variables across the dataset.

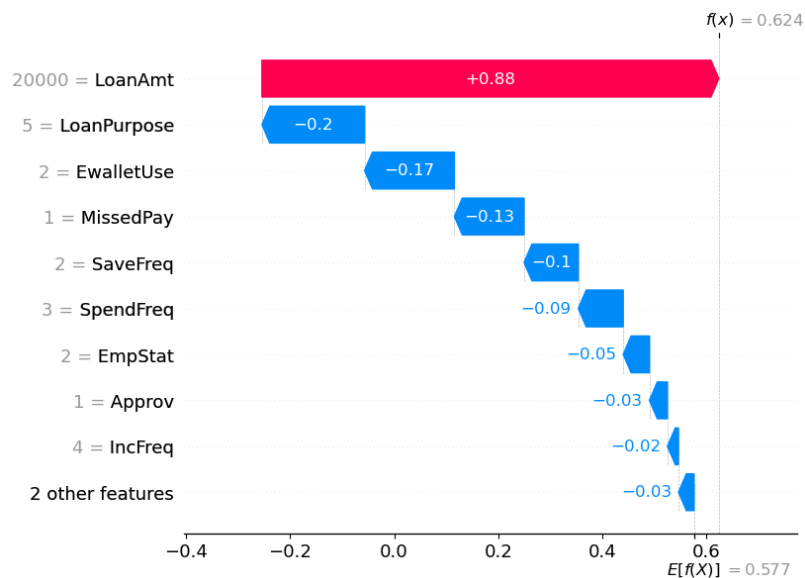


Figure 3. SHAP Waterfall Plot

Table 3 presents the confusion matrix for the selected CatBoost model on the 154-observation test set. The model correctly classified 107 observations, yielding an accuracy of

69.48%. When default is treated as the positive class, the model correctly identified 43 borrowers who did not fully repay their loans, while 17 actual defaulters were classified as repaid. In addition, 30 borrowers who fully repaid their loans were incorrectly classified as defaulters, whereas 64 fully repaid loans were correctly identified. Overall, these results show that the model can distinguish between repayment outcomes, but misclassification remains a consideration for potential credit-risk applications.

Table 3. Confusion Matrix of the CatBoost Model

Actual outcome	Predicted Class 0 (Default)	Predicted Class 1 (Repaid)	Total
Default (Class 0)	43 (TP)	17 (FN)	60
Repaid (Class 1)	30 (FP)	64 (TN)	94
Total	73	81	154

Note: Default (Class 0) is the positive class for default-focused metrics. TP = true positive; FN = false negative; FP = false positive; TN = true negative.

Next, the table below presents the performance estimates and corresponding 95% confidence intervals (CIs) for CatBoost. The model achieved an accuracy of 0.70 and an AUC-ROC of 0.81, indicating its ability to discriminate between the two repayment outcomes in the test sample. The CIs also show the variation in the reported performance estimates.

Table 4. CatBoost Performance Summary

Metric	Reported result	95% CI
Accuracy	0.70	0.615–0.760
Precision	0.79	0.670–0.830
Recall	0.68	0.610–0.780
F1-score	0.73	0.650–0.800
AUC-ROC	0.81	0.730–0.870

Note: Precision, recall, and F1-score are reported for Class 1 (fully repaid) as the positive class using binary classification metrics. Accuracy is the proportion of correctly classified observations, and AUC-ROC measures the model's ability to discriminate between the two classes. The reported confidence intervals correspond to the stated performance estimates.

Furthermore, the table below presents the calibration analysis of CatBoost, based on 154 test observations, with mean predicted default probabilities generally close to the observed default rates across the five probability intervals. First, the 20–40% interval showed a mean predicted probability of 30% and an observed default rate of 29%. By contrast, the 80–100% interval showed corresponding values of 89% and 88%. Meanwhile, the 60–80% interval recorded a slightly higher observed default rate (73%) than its mean predicted probability (70%). Overall, these results suggest reasonable agreement between predicted probabilities and observed default frequencies within the reported intervals.

Table 5. Calibration Analysis of CatBoost

Predicted default probability	Sample count	Mean predicted probability	Observed default rate
0–20%	30	0.12	0.1
20–40%	35	0.3	0.29
40–60%	34	0.5	0.47
60–80%	30	0.7	0.73
80–100%	25	0.89	0.88
Total	154	-	-

Furthermore, the performance of CatBoost across five cross-validation folds was examined, as presented in the table below. The reported fold-level results indicated that the model achieved a mean accuracy of 0.70, precision of 0.79, recall of 0.68, F1-score of 0.73, and AUC-ROC of 0.81. These results suggest that accuracy, precision, and recall exhibited greater variation than the F1-score and AUC-ROC, indicating performance consistency within the cross-validation procedure.

Table 6. CatBoost Performance Across Five Cross-Validation Folds

Validation fold	Accuracy	Precision	Recall	F1-score	AUC-ROC
Fold 1	0.65	0.76	0.62	0.71	0.78
Fold 2	0.72	0.82	0.71	0.76	0.84
Fold 3	0.68	0.79	0.65	0.73	0.81
Fold 4	0.75	0.85	0.73	0.74	0.83
Fold 5	0.70	0.74	0.69	0.72	0.79
Mean	0.70	0.79	0.68	0.73	0.81
Standard deviation	0.02	0.03	0.03	0.02	0.02

Next, the table below shows the CatBoost model's classification performance when default is treated as the positive class. The model achieved a default precision of 0.589, meaning that 58.9% of borrowers classified as defaulters were actual defaulters in the test set. Default recall was 0.717, indicating that the model correctly identified 71.7% of actual defaulters. Specificity was 0.681, meaning that 68.1% of borrowers who fully repaid their loans were correctly identified as non-defaulters. Furthermore, the default F1-score was 0.647, and overall accuracy was 0.695, indicating that the model identified most actual defaulters but also misclassified some borrowers in both directions.

Table 7. Default Detection Performance of CatBoost

Default-focused Metric	Calculation	Result
Default precision	$(43/(43+30))$	0.589
Default recall (sensitivity)	$(43/(43+17))$	0.717
Specificity	$(64/(64+30))$	0.681
Default F1-score	$(2(0.589)(0.717)/(0.589+0.717))$	0.647
Accuracy	$(107/154)$	0.695

Note: The default (Class 0) is treated as the positive class. Precision, recall (sensitivity), and F1-score are class-specific binary metrics for default detection. Specificity measures the proportion of fully repaid loans that are correctly classified as non-defaults.

CONCLUSIONS

This study examined the predictive performance of ML in classifying loan repayment outcomes among 512 borrowers using bootstrapped digital lending platforms in Metro Manila, Philippines. Four models were compared using classification performance metrics, including XGBoost, LightGBM, Logistic Regression, and CatBoost. Among these models, CatBoost achieved the highest reported AUC-ROC of approximately 0.81, indicating the strongest discriminatory performance among the evaluated models under the study's evaluation procedure, with an accuracy of 0.70. This result suggests that the model could distinguish between repayment outcomes, though with moderate predictive performance. Furthermore, these results support the potential applicability of gradient boosting methods to credit-risk prediction in categorical-heavy lending datasets, although the findings do not establish that the model will perform equally well in other borrower populations or lending environments.

Next, feature importance analysis revealed that loan characteristics accounted for

approximately 90.81% of the model's reported feature importance, with loan amount contributing approximately 82.85%. Furthermore, SHAP analysis showed that higher loan amounts were generally associated with higher predicted default risk in the observed dataset, suggesting that loan amount is the dominant predictor used by CatBoost in generating its predictions. However, feature importance and SHAP values represent model-based contributions rather than causal effects. For financial capacity indicators and behavioral traits, both accounted for approximately 6.28% and 2.91%, respectively, of the reported feature importance, suggesting that, relative to loan characteristics, these provide less predictive information for CatBoost, the selected model, in this particular dataset.

Theoretical Implications

This study contributes to the conceptual understanding of loan default prediction by examining three predictor domains: loan characteristics, financial capacity indicators, and behavioral traits. Rather than testing a single established theory, the study derives its predictor structure from a synthesis of prior credit-risk and ML literature and evaluates the relative predictive contributions of these domains within a specific digital lending context. Furthermore, the findings suggest that the predictive usefulness of these domains is not necessarily equivalent. In the present dataset, loan characteristics accounted for the largest share of the selected model's reported feature importance. Meanwhile, financial capacity indicators and behavioral traits contributed comparatively less, providing only an empirical basis for further examining whether the relative predictive contributions of conventional loan-related variables and borrower-level characteristics differ across lending environments.

Furthermore, this study illustrates the distinction between conceptual relevance and predictive contribution. Although financial capacity and behavioral traits are relevant domains in the broader credit-risk literature, their comparatively small feature importance scores indicate that their contribution to the selected model's predictions was limited in this sample. Hence, this does not establish that these domains lack theoretical relevance or are unrelated to repayment outcomes. Instead, it highlights the importance of empirically evaluating the predictive contribution of proposed variables within the population and modeling context under investigation.

Finally, the study demonstrates how explainability techniques can be incorporated into gradient-boosting-based credit-risk prediction. By combining model-level feature importance with SHAP explanations, it provides a more detailed account of how the model uses borrower and loan information to generate predictions. Hence, this paper and its results contribute methodologically to applied credit-risk research by showing how predictive performance and model interpretability can be examined together.

Recommendations

Based on the findings, bootstrapped digital lending platforms may consider evaluating gradient boosting models, particularly CatBoost, as supplementary tools for credit risk assessment. The model's performance on the present dataset suggests potential usefulness in distinguishing borrowers with different observed repayment outcomes. However, its predictions should not be treated as definitive determinations of an individual's creditworthiness, and its operational suitability should be established through additional validation before deployment.

First, loan amount emerged as the most influential predictor in the selected model. Therefore, lending platforms may examine how loan amount, alongside other relevant borrower and loan characteristics, contributes to predicted repayment risk within their own customer portfolios. This is because such assessments should not rely on loan amount alone. Integrating feature importance

with SHAP can help credit-risk analysts better understand the factors that contributed to the model's predictions, especially the propensity of users to default on loans. These explanations could support the review of individual cases and the identification of patterns requiring further assessment. Still, model explanations must be considered alongside independently verified information, established lending criteria, and appropriate human review for credit decisions.

Before operational implementation, lending platforms must evaluate the model using independently collected, representative data that include borrowers from locations and lending environments beyond Metro Manila. Further assessments should examine performance stability, probability calibration, false-positive and false-negative rates, and the potential for unequal predictive errors across borrower groups. Where possible, future validation should use verified repayment records rather than relying exclusively on self-reported survey responses.

LIMITATION & FURTHER RESEARCH

This study demonstrated the value of combining predictive modeling with model-interpretation techniques, where feature importance summarized the relative contributions of the predictors, while SHAP provided additional insight into their contributions to individual predictions and the overall direction of the model's predictions. Hence, these techniques improve the interpretability of the selected model. These findings can contribute to the credit-risk prediction literature by examining a comparatively less-studied setting: bootstrapped digital lending platforms in the Philippines, and the results can provide preliminary empirical evidence on how loan characteristics, financial capacity indicators, and behavioral traits.

These variables can also contribute to machine-learning predictions of repayment outcomes in this setting. Nevertheless, the study's self-reported data, geographic restriction to Metro Manila, and reliance on the observed sample limit the generalizability of its conclusions. Future research may examine additional borrower populations, independently verified repayment records, and alternative validation settings to determine whether the observed predictive patterns remain consistent. Lastly, future research may extend this study by examining additional borrower characteristics, alternative measures of financial behavior, and repayment records collected over time.

REFERENCES

- Aattouf, A., El Kafhali, S., & Saadi, Y. (2021). Predicting the e-signing likelihood of loan using machine learning models combining clustering with supervised learning. In *Advances on smart and soft computing: Proceedings of ICACIn 2021* (pp. 121–131). Springer.
- Adamu, I. Y. (2018). Idea of collateral and guarantor in Islamic bank financing. *SEISENSE Journal of Management*, 1(5), 49–57. <https://doi.org/10.5281/zenodo.1474664>
- Adiani, W. (2021). External drivers of e-wallet continuance and loyalty among mid-to-late adulthood consumers in Indonesia. *Applied Quantitative Analysis*, 1(1), 35–54. <https://doi.org/10.31098/quant.3991>
- Anand, M., Velu, A., & Whig, P. (2022). Prediction of loan behaviour with machine learning models for secure banking. *Journal of Computer Science and Engineering (JCSE)*, 3(1), 1–13.
- Arowolo, M. O., Adeniyi, O. F., Adebisi, M. O., & Ogundokun, R. O. (2022). A prediction model for bank loans using agglomerative hierarchical clustering with classification approach. *Covenant Journal of Informatics and Communication Technology*, 10(2). <https://journals.covenantuniversity.edu.ng/index.php/cjict/article/view/3147>
- Bastani, H., Ascarza, E., & Chod, J. (2019). *Improving credit scoring with machine learning* (Working Paper No. 20-006). Harvard Business School.
- Chen, M., Wu, Q., & Chen, H. (2021). A deep learning framework for credit risk assessment in P2P

- lending. *Electronic Commerce Research and Applications*, 48, Article 101073. <https://doi.org/10.1016/j.elerap.2021.101073>
- Dansana, D., Patro, S. G. K., Mishra, B. K., Prasad, V., Razak, A., & Wodajo, A. W. (2024). Analyzing the impact of loan features on bank loan prediction using random forest algorithm. *Engineering Reports*, 6(2), Article e12707.
- Dastile, X., Celik, T., & Potsane, M. (2020). Statistical and machine learning models in credit scoring: A systematic literature survey. *Applied Soft Computing*, 91, Article 106263. <https://doi.org/10.1016/j.asoc.2020.106263>
- Dhruv, C., Paul, D., Kumar, M. H., & Reddy, M. S. (2023). Framework for bank loan re-payment prediction and income prediction. In *2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC)* (pp. 833–840). IEEE.
- Dutta, P. (2021). A study on machine learning algorithm for enhancement of loan prediction. *International Research Journal of Modernization in Engineering Technology and Science*. <https://doi.org/10.6084/M9.JETIR.JETIRDY06108>
- Faraji, Z. (2022). A review of machine learning applications for credit card fraud detection with a case study. *SEISENSE Journal of Management*, 5(1), 49–59. <https://doi.org/10.33215/sjom.v5i1.770>
- Harjanto, N., & Nurim, Y. (2023). The moderating role of liquidity on the relationship between rural banks' human capital quality and credit risk. *Review of Integrative Business and Economics Research*, 12(3), 171–181.
- Hegde, P., Shetty, C., Rashmi, N., & Dhananjaya, B. (2022). Predictive analysis of loan data using machine learning. In *2022 International Conference on Artificial Intelligence and Data Engineering (AIDE)* (pp. 272–276). IEEE.
- Jagtiani, J., & Lemieux, C. (2019). The roles of alternative data and machine learning in fintech lending: Evidence from the LendingClub consumer platform. *Journal of Economics and Business*, 100, 105–121. <https://doi.org/10.1016/j.jeconbus.2018.11.002>
- Jian Ai, Y., Chew Sze, C., Sook Fern, Y., Hen Toong, T., & Boon Chian, C. (2021). The use of e-wallet among Gen-Y in Malaysia during the global pandemic: Using PLS-SEM analysis. *Applied Quantitative Analysis*, 1(1), 9–17. <https://doi.org/10.31098/quant.597>
- Kokate, S., & Chetty, M. S. R. (2021). Credit risk assessment of loan defaulters in commercial banks using voting classifier ensemble learner machine learning model. *International Journal of Safety and Security Engineering*, 11(5), 565–572.
- Kumar, A., Dugyala, R., & Bhattacharya, P. (2021). Prediction of loan scoring strategies using deep learning algorithm for banking system. In P. K. Singh, Z. Polkowski, S. Tanwar, S. K. Pandey, G. Matei, & D. Pirvu (Eds.), *Innovations in information and communication technologies (IICT-2020)*. Springer. https://doi.org/10.1007/978-3-030-66218-9_13
- Kwofie, C., Owusu-Ansah, C., & Boadi, C. (2015). Predicting the probability of loan-default: An application of binary logistic regression. *Research Journal of Mathematics and Statistics*, 7(4), 46–52.
- Lee, J., & Kim, Y. (2020). Fintech credit scoring and borrower behavior: Are alternative data useful? *Finance Research Letters*, 36, Article 101606. <https://doi.org/10.1016/j.frl.2019.101606>
- Madaan, M., Kumar, A., Keshri, C., Jain, R., & Nagrath, P. (2021). Loan default prediction using decision trees and random forest: A comparative study. *IOP Conference Series: Materials Science and Engineering*, 1022(1), Article 012042. <https://doi.org/10.1088/1757-899X/1022/1/012042>
- Manglani, R., & Bokhare, A. (2021). Logistic regression model for loan prediction: A machine learning approach. In *2021 Emerging Trends in Industry 4.0 (ETI 4.0)* (pp. 1–6). IEEE.
- Ndayishimiye, A., & Kim, J. (2021). Predicting default risk using machine learning algorithms in

- microfinance. *Journal of Risk Finance*, 22(3), 235–253.
- Patiu, L. S., & Eleazar, V. (2024). Bank-specific, industry-specific, and macroeconomic determinants of credit risk: Empirical evidence from the ASEAN5 banking sector. *Review of Integrative Business and Economics Research*, 13(4), 436–454.
- Prasanth, C., Kumar, R. P., Rangesh, A., Sasmitha, N., & Dhiyanesh, B. (2023). Intelligent loan eligibility and approval system based on random forest algorithm using machine learning. In *2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)* (pp. 84–88). IEEE.
- Purnawan, V., & Danu Prasetyo, A. (2021). Domestic and foreign banks' stability in Indonesia: The grey zone trap and key determinants. *Applied Quantitative Analysis*, 1(2), 1–16. <https://doi.org/10.31098/quant.596>
- Serrano-Cinca, C., Gutiérrez-Nieto, B., & López-Palacios, L. (2015). Determinants of default in P2P lending. *PLOS ONE*, 10(10), Article e0139427. <https://doi.org/10.1371/journal.pone.0139427>
- Sheikh, M. A., Goel, A. K., & Kumar, T. (2020). An approach for prediction of loan approval using machine learning algorithm. In *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 490–494). IEEE.
- Shinde, A. B., Patil, Y., Kotian, I., Shinde, A., & Gulwani, R. R. (2022). Loan prediction system using machine learning. *ITM Web of Conferences*, 44, Article 03019. <https://doi.org/10.1051/itmconf/20224403019>
- Statista. (n.d.). *Number of fintech firms in the Philippines*. <https://www.statista.com/statistics/1247613/number-of-fintech-firms-philippines/>
- Tonik Bank. (2024). *Loan statistics in the Philippines 2023*. <https://tonikbank.com/blog/loan-statistics-philippines-2023>
- Zhang, Y., Zhang, X., & Tang, Q. (2022). Machine learning approaches to predict consumer credit default: Evidence from peer-to-peer lending platforms. *Expert Systems with Applications*, 190, Article 116236.