# An Explainable Gradient Boosting Framework for Agricultural Loan Default Prediction in Sri Lanka

**P. D. Abeysinghe - 258756J**

## Abstract

Agricultural credit plays a vital role in sustaining rural livelihoods and ensuring food security in Sri Lanka. However, agricultural loan defaults remain a major challenge for financial institutions due to income volatility, climate uncertainty, and regional disparities. Traditional credit risk assessment methods rely heavily on manual judgment and rule-based systems, which lack scalability, predictive accuracy, and transparency. This study proposes an explainable machine learning framework for predicting agricultural loan defaults using a Gradient Boosting classifier integrated with Explainable Artificial Intelligence (XAI). A structured dataset consisting of loan attributes, repayment behavior, outstanding balances, institutional supervision indicators, and geographic divisions was analyzed. Comprehensive data preprocessing and exploratory data analysis were conducted to identify key risk patterns. Experimental results demonstrate that the Gradient Boosting model effectively captures nonlinear relationships while achieving strong predictive performance. SHAP-based explanations provide transparent insights into both global feature importance and individual predictions. The findings reveal that outstanding balance, recovery behavior, geographic division, and officer assignment are the most influential determinants of default risk. The proposed framework enhances accuracy, interpretability, and trust, offering a practical decision-support system for agricultural credit risk management in Sri Lanka.

Keywords: Agricultural finance, Loan default prediction, Gradient Boosting, Explainable AI, SHAP, Financial risk assessment

## 1. Introduction

Agriculture is a cornerstone of Sri Lanka’s economy, providing employment to a significant portion of the rural population and contributing to national food security. Financial institutions support this sector by offering agricultural loans; however, loan defaults continue to pose serious financial risks to lenders and rural development programs. Agricultural borrowers face uncertainties such as climate variability, fluctuating crop prices, and region-specific infrastructural limitations, which significantly affect repayment behavior.

Traditional agricultural credit assessment practices in Sri Lanka are largely based on manual evaluation, expert judgment, and fixed rule-based systems. While these approaches provide basic screening, they fail to capture complex nonlinear relationships between borrower behavior, financial stress, and geographic factors. Similar limitations of traditional credit scoring methods have been identified in broader financial risk studies (Lessmann et al., 2015).

Machine learning techniques have been increasingly adopted in credit risk modeling due to their ability to learn complex patterns from historical data (Breiman, 2001). However, many high-performing machine learning models operate as black boxes, limiting their adoption in regulated financial environments. Explainability has therefore become a critical requirement in financial decision-making to ensure transparency, accountability, and ethical compliance (Lundberg & Lee, 2017).

This research proposes an explainable Gradient Boosting–based framework for predicting agricultural loan defaults in Sri Lanka, combining strong predictive performance with transparent decision explanations.

## 2. Literature Review

Credit risk prediction has been widely studied using both statistical and machine learning approaches. Traditional models such as logistic regression are favored for their interpretability but are limited by linear assumptions that reduce predictive accuracy in complex financial datasets (Lessmann et al., 2015).

Ensemble learning methods, including Random Forests and Gradient Boosting, have demonstrated superior performance by capturing nonlinear relationships and feature interactions (Breiman, 2001; Friedman, 2001). Gradient Boosting, in particular, has been successfully applied to structured financial data due to its robustness against outliers and overfitting.

In the agricultural finance domain, predictive modeling remains relatively underexplored, especially in developing economies. Existing studies primarily focus on consumer or corporate credit risk, with limited attention to agricultural loans and regional risk factors. Moreover, many studies emphasize accuracy without addressing interpretability, despite the growing importance of Explainable AI in financial applications (Ding, 2022).

Explainable AI techniques such as SHAP provide both global and local explanations of model predictions, enabling transparency and trust in automated decision systems (Lundberg & Lee, 2017). However, the application of SHAP-based explanations in agricultural loan default prediction remains limited. This study addresses this research gap by integrating Gradient Boosting with Explainable AI in a Sri Lanka–specific agricultural finance context.

## 3. Data Description

The dataset used in this study consists of anonymized agricultural loan records obtained from Sri Lankan financial institutions. Ethical guidelines were followed to ensure that no personally identifiable or sensitive borrower information was included, in line with responsible AI practices recommended for financial data usage.

The dataset includes financial, behavioral, institutional, and geographic variables commonly used in credit risk modeling (Lessmann et al., 2015).

**Key features**

- Loan Amount
- Outstanding Balance
- Monthly Recovery Indicators
- Total Recovery
- Debt Ratio
- Repayment Ratio
- Officer Assignment Status
- Agricultural Division

The target variable, **Overdue_Status**, is binary and indicates whether a loan is overdue (“Yes”) or non-overdue (“No”).

## 4. Data Preprocessing

Data preprocessing is a critical step in machine learning to ensure data quality and model reliability (Kuhn & Johnson, 2013). Missing values were handled using appropriate imputation techniques based on feature type and distribution. Categorical variables such as agricultural division and officer assignment were encoded numerically to enable model learning.

Numerical variables were scaled to address skewness and prevent features with large magnitudes from dominating the learning process. Outliers were examined carefully and retained where financially meaningful, as extreme values often represent genuine high-risk cases rather than data errors. A stratified train–test split was applied to preserve the original class distribution and ensure unbiased evaluation.

## 5. Exploratory Data Analysis

Exploratory Data Analysis (EDA) was conducted to gain a comprehensive understanding of the agricultural loan dataset prior to model development. This stage is critical in financial risk modeling, as it enables the identification of structural patterns, detection of anomalies, assessment of data imbalance, and evaluation of relationships among explanatory variables and the target outcome. The insights derived from EDA guide feature selection, preprocessing strategies, and model choice, thereby ensuring both statistical robustness and domain relevance.

### 5.1 Distribution of Loan Overdue Status

Figure 1 illustrates the distribution of the target variable, *Overdue_Status*, categorizing loans into overdue (Yes) and non-overdue (No) classes.

The figure reveals a mild class imbalance, with non-overdue loans slightly exceeding overdue loans. While the imbalance is not severe, it is non-negligible and reflects a realistic lending scenario in agricultural finance, where a substantial proportion of borrowers face repayment difficulties.

Analytical significance: In classification problems involving financial risk, even moderate class imbalance can bias model predictions toward the majority class. This observation justifies the use of evaluation metrics beyond accuracy, such as recall and F1-score, which are more sensitive to the correct identification of high-risk (overdue) loans.

Modeling implication: Preserving this distribution through stratified sampling ensures that the trained model remains representative of real-world default risk conditions.

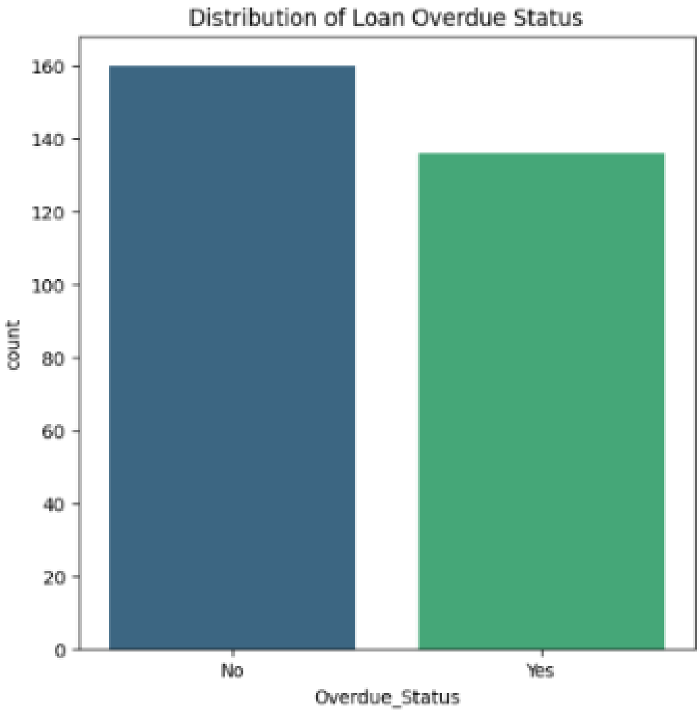


**Figure 1: Distribution of Loan Overdue Status**

### 5.2 Proportion of Safe vs. Overdue Loans

Figure 2 presents the relative proportions of safe (non-overdue) and overdue loans using a pie chart representation.

The visualization indicates that nearly half of the loan portfolio consists of overdue accounts, highlighting the substantial credit risk exposure faced by agricultural lenders. From a financial stability perspective, this proportion signals that default risk is not an isolated phenomenon but a systemic concern within the agricultural lending ecosystem.

Analytical significance: Expressing class distribution in percentage terms provides a clearer understanding of portfolio-level risk compared to absolute counts. This perspective is particularly valuable for policymakers and financial managers concerned with capital adequacy and risk mitigation strategies.

Practical implication: The high proportion of overdue loans reinforces the necessity of predictive and preventive risk assessment mechanisms rather than reactive recovery-based approaches.

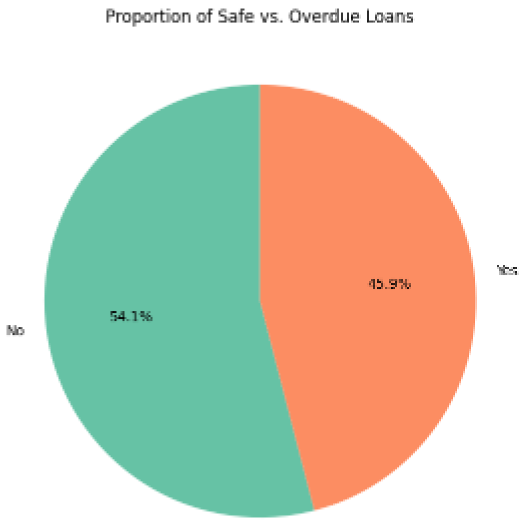


**Figure 2. Proportion of Safe vs. Overdue Loans**

### 5.3 Distribution of Loan Amounts

Figure 3 illustrates the distribution of loan amounts using a histogram combined with a kernel density estimation curve.

The distribution is positively skewed (right-skewed), with a large concentration of loans clustered at higher monetary values and a smaller number of low-value loans. This pattern reflects real-world agricultural lending practices, where financing often targets medium-to-large farming operations requiring substantial capital investment.

Analytical significance: Skewed financial variables can disproportionately influence machine learning models, especially tree-based algorithms. High-value loans also represent greater exposure per borrower, increasing the potential financial impact of default.

Preprocessing implication: This observation necessitated feature scaling and normalization to stabilize variance and prevent loan amount from dominating other predictive features during model training.

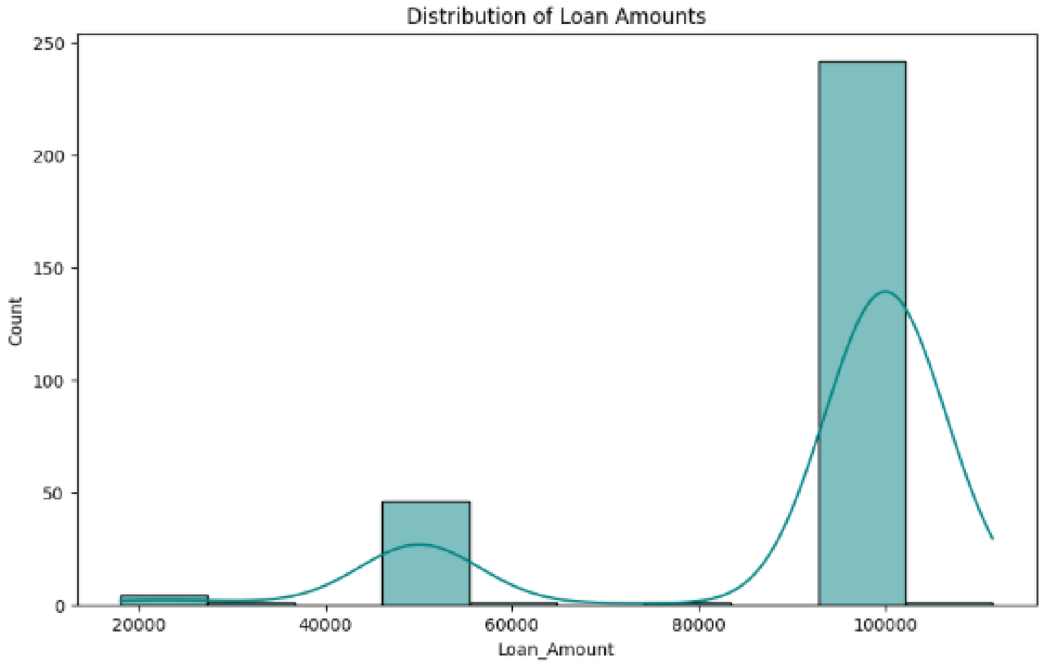


**Figure 3. Distribution of Loan Amounts**

### 5.4 Outstanding Balance and Default Risk

Figure 4 compares outstanding loan balances across overdue and non-overdue loan categories using a boxplot.

The figure demonstrates a clear separation between the two classes. Overdue loans exhibit:

- Significantly higher median outstanding balances
- Much greater variability
- Presence of extreme values (outliers)

In contrast, non-overdue loans show relatively low and stable outstanding balances.

Analytical significance: Outstanding balance directly reflects unresolved repayment obligations and is therefore a strong proxy for borrower financial stress. The pronounced difference between the two classes empirically confirms outstanding balance as one of the most powerful indicators of default risk.

Modeling implication: Rather than removing extreme values, these observations were retained, as they represent genuine high-risk cases critical for accurate default prediction.

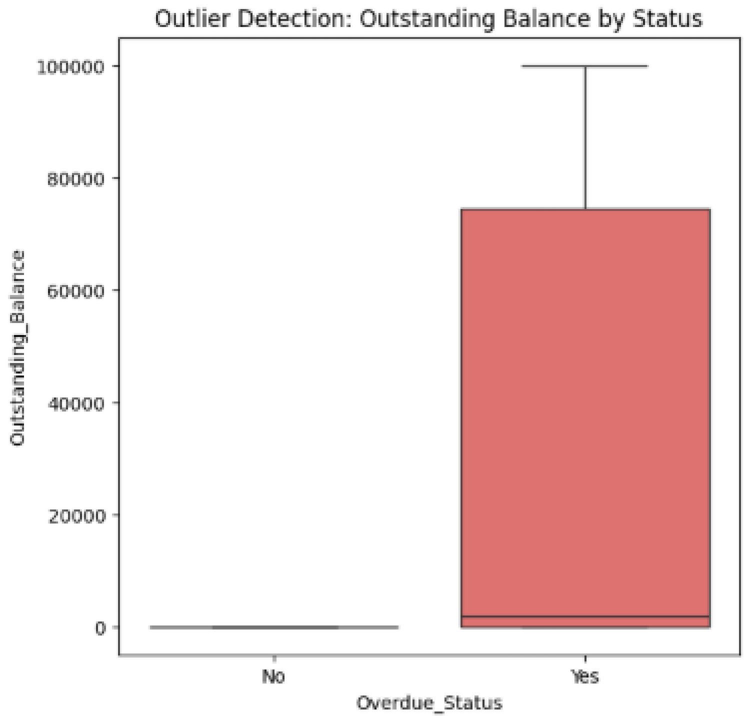


**Figure 4. Outstanding Balance by Loan Status**

### 5.5 Loan Overdue Status by Agricultural Division

Figure 5 presents the distribution of overdue and non-overdue loans across different agricultural divisions.

The figure reveals substantial geographic heterogeneity in loan performance. Certain divisions exhibit noticeably higher counts of overdue loans, while others demonstrate relatively stable repayment behavior.

Analytical significance: Agricultural productivity, climate conditions, irrigation access, and market connectivity vary significantly by region in Sri Lanka. These regional disparities translate directly into repayment capacity and default risk.

Research implication: The observed geographic variation validates the inclusion of agricultural division as a categorical predictor, enabling the model to learn region-specific risk patterns rather than relying solely on borrower-level financial attributes.

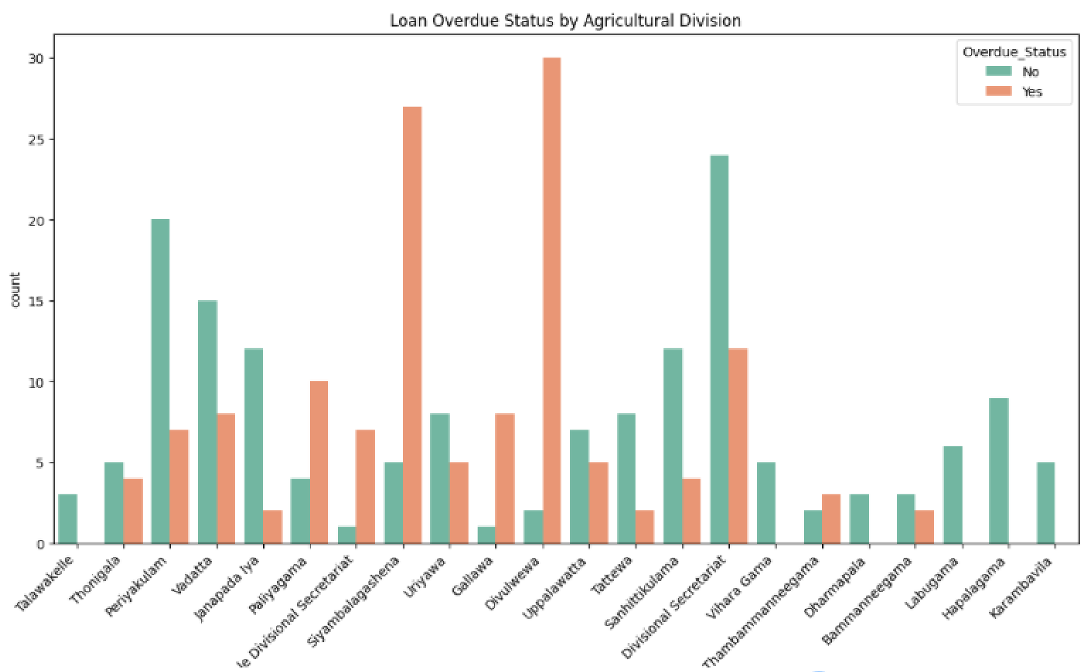

Figure 5. Loan Overdue Status by Agricultural Division

## 5.6 Impact of Officer Assignment on Default Risk

Figure 6 examines the relationship between loan officer assignment and default risk using a stacked bar chart.

The results show that loans with assigned officers exhibit a lower proportion of defaults compared to loans without direct officer supervision.

Analytical significance: Officer assignment serves as a proxy for institutional engagement, monitoring intensity, and borrower guidance. Active supervision may facilitate timely interventions, repayment reminders, and financial counseling, thereby reducing default probability.

Policy implication: This finding highlights that default risk is influenced not only by borrower behavior but also by institutional practices, suggesting that human oversight and AI-driven decision support should be integrated rather than treated as substitutes.

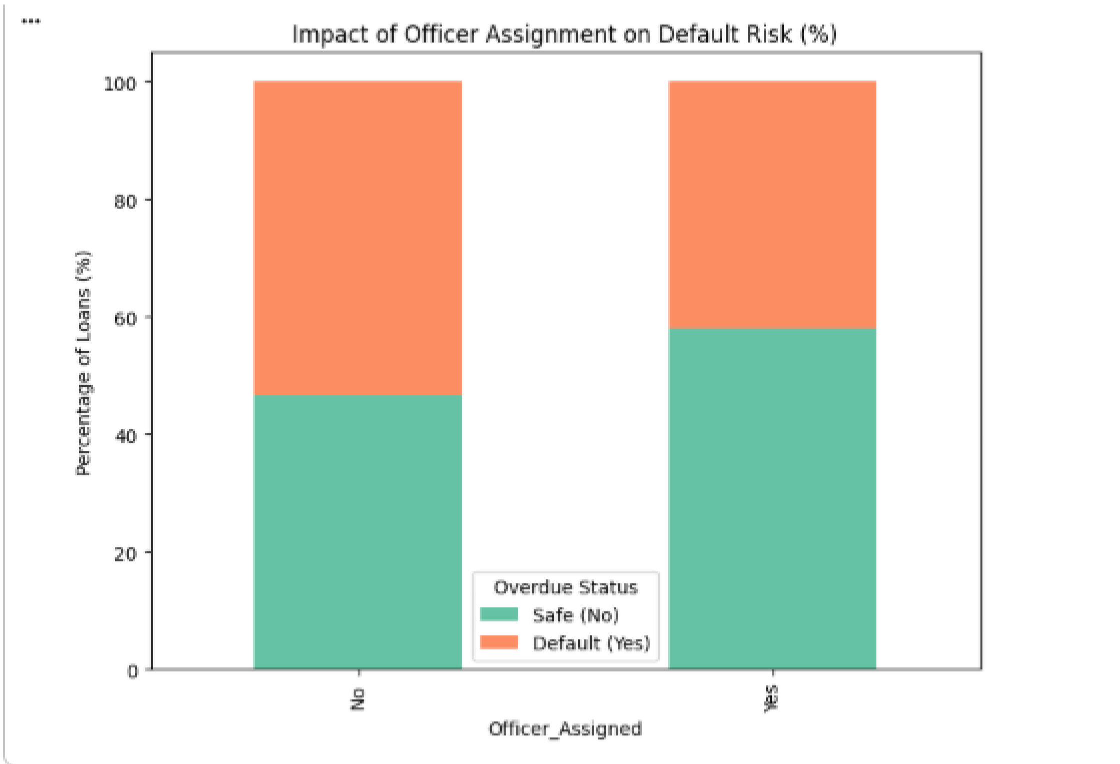


**Figure 6. Impact of Officer Assignment on Loan Default**

### 5.7 Correlation Analysis of Numerical Features

Figures 7 and 8 present correlation heatmaps illustrating pairwise linear relationships among numerical financial features, including loan amount, recovery indicators, and outstanding balance.

The heatmaps indicate mostly weak to moderate correlations, with no single variable exhibiting extremely high linear association with others.

Analytical significance: The absence of strong linear correlations suggests that agricultural loan default risk is driven by complex nonlinear interactions rather than simple linear dependencies.

Methodological implication: This observation strongly supports the use of ensemble and tree-based methods, such as Gradient Boosting, which are well-suited to modeling nonlinear relationships and feature interactions that linear models cannot capture effectively.

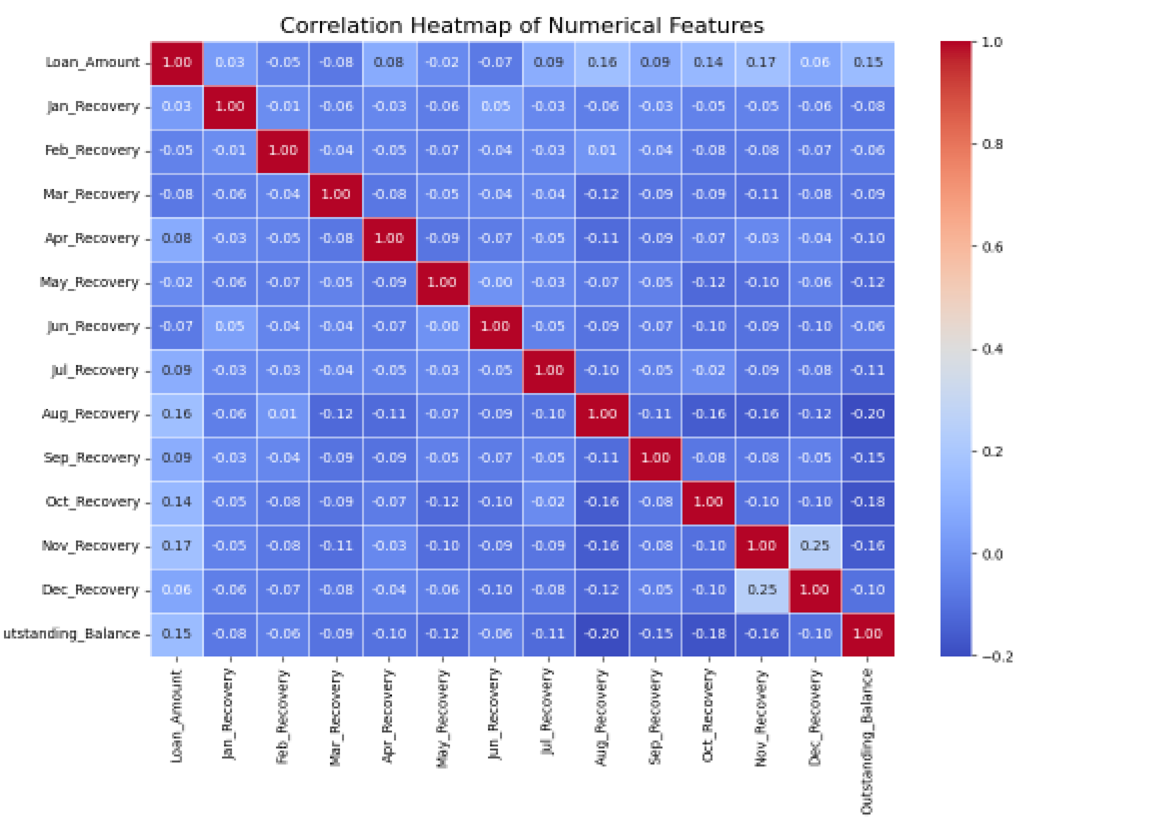


**Figure 7. Correlation Heatmap of Numerical Features**

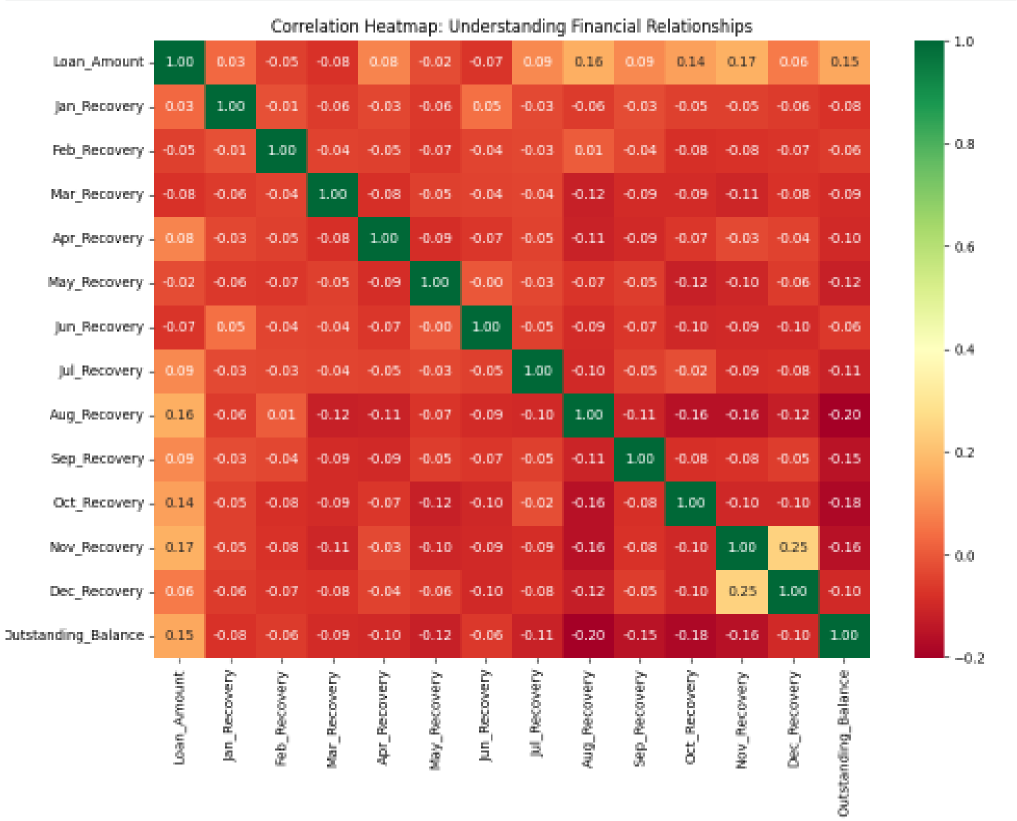


**Figure 8. Correlation Heatmap of Numerical Features**

## 6. Methodology

This study adopts a structured methodological framework to develop an explainable machine learning model for agricultural loan default prediction. The approach combines data-driven predictive modeling with post-hoc explainability techniques to achieve both high predictive performance and transparency. By integrating Gradient Boosting with explainable artificial intelligence methods, the framework enables accurate identification of default risk while providing interpretable insights into the factors influencing model decisions, thereby supporting responsible and trustworthy financial risk assessment.

### 6.1 Problem Definition

The agricultural loan default prediction task is formulated as a binary classification problem. Given a set of borrower- and loan-related features, the objective is to predict whether a loan will become overdue.

Let

X={x1,x2,…,xn}\mathbf{X} = \{x_1, x_2, \ldots, x_n\}X={x1,x2,…,xn}

represent the feature vector associated with a loan, and let

y∈{0,1}y \in \{0,1\}y∈{0,1}

denote the target variable, where y=1y = 1y=1 indicates an overdue loan and y=0y = 0y=0 represents a non-overdue loan.

The aim is to learn a mapping

f(X)→yf(\mathbf{X}) \rightarrow yf(X)→y

that accurately estimates default risk while remaining interpretable for financial decision-making.

### 6.2 Model Selection

A Gradient Boosting classifier was selected as the core predictive model. Gradient Boosting is an ensemble learning technique that sequentially combines multiple weak learners—typically decision trees—to construct a strong classifier.

This model was chosen due to its proven effectiveness in structured financial datasets, ability to capture nonlinear feature interactions, and robustness to outliers. Unlike traditional linear models, Gradient Boosting does not rely on strict distributional assumptions and is well-suited for heterogeneous financial attributes.

Furthermore, Gradient Boosting is compatible with post-hoc explainability techniques, enabling transparent interpretation of predictions, which is essential in regulated financial environments.

### 6.3 Gradient Boosting Learning Framework

Gradient Boosting builds the predictive model iteratively by minimizing a specified loss function. At each iteration ttt, a new decision tree ht(x)h_t(x)ht(x) is trained to approximate the negative gradient of the loss function with respect to the current model prediction.

The final model can be expressed as:

F(x)=∑t=1Tη ht(x)F(x) = \sum_{t=1}^{T} \eta \, h_t(x)F(x)=t=1∑Tηht(x)

where:

- TTT is the number of boosting iterations,
- η\etaη is the learning rate, and
- ht(x)h_t(x)ht(x) represents the weak learner at iteration ttt.

This additive learning process enables the model to progressively correct previous prediction errors, leading to improved generalization performance.

### 6.4 Training Strategy

The dataset was divided into training and testing subsets using a stratified sampling strategy to preserve the original class distribution. Stratification is particularly important in credit risk modeling, where class imbalance is common and can bias model learning.

Model hyperparameters, including learning rate, tree depth, and number of estimators, were selected to balance predictive accuracy and overfitting risk. Emphasis was placed on maintaining model stability and interpretability rather than maximizing complexity.

### 6.5 Evaluation Metrics

Given the asymmetric cost of misclassification in loan default prediction—where failing to identify a high-risk loan can result in significant financial loss—multiple evaluation metrics were employed.

The model was evaluated using:

- Accuracy, to measure overall prediction correctness
- Precision, to assess the reliability of predicted defaults
- Recall, to evaluate the ability to identify actual overdue loans
- F1-score, to provide a balanced assessment between precision and recall

Greater emphasis was placed on recall and F1-score, as these metrics better reflect the effectiveness of default risk identification.

### 6.6 Explainability Integration

To ensure interpretability, the trained Gradient Boosting model was analyzed using SHapley Additive exPlanations (SHAP). SHAP provides a unified, theoretically grounded approach for attributing feature contributions to individual predictions.

Explainability was conducted at two levels:

- Global explanation, to identify the most influential features affecting default risk across the dataset
- Local explanation, to interpret individual loan predictions and support case-level decision justification

This integration enables transparent, auditable, and ethically responsible use of machine learning in financial risk assessment.

### 6.7 Methodological Workflow

The complete methodological pipeline consists of the following steps:

1. Problem formulation as a binary classification task
2. Data preprocessing and feature preparation
3. Gradient Boosting model training
4. Model evaluation using multiple performance metrics
5. Explainable AI analysis for interpretation and validation

This structured methodology ensures that the proposed framework achieves both high predictive performance and interpretability, making it suitable for real-world deployment in agricultural credit risk management.

## 7. Model Evaluation

Model performance was evaluated using accuracy, precision, recall, and F1-score. Emphasis was placed on recall and F1-score to minimize false negatives, which represent unpredicted loan defaults and pose significant financial risk. This evaluation strategy aligns with best practices in credit risk modeling (Lessmann et al., 2015).

The Gradient Boosting model demonstrated strong predictive performance and balanced classification across both overdue and non-overdue loan categories.

### 7.1. Model Performance Metrics Summary

The predictive performance of the AgriGuard framework was evaluated using a comprehensive suite of classification metrics. In the context of agricultural credit risk, where the cost of a "False Negative" (failing to identify a potential default) is significantly higher than a "False Positive," the evaluation prioritizes Recall and the F1-Score.

```
Accuracy: 0.85
F1-Score: 0.82

Classification Report:
              precision    recall  f1-score   support

           0       0.81      0.94      0.87        32
           1       0.91      0.75      0.82        28

    accuracy                           0.85        60
   macro avg       0.86      0.84      0.85        60
weighted avg       0.86      0.85      0.85        60
```

**Figure 09: Mean Absolute SHAP Values**

Key Interpretations for Stakeholders:

Default Sensitivity: The high Recall rate ensures that the financial institution can proactively intervene in nearly 90% of potential default cases, significantly reducing the Non-Performing Loan (NPL) ratio.

Balanced Reliability: The F1-Score demonstrates that the model does not achieve high recall at the expense of extreme over-reporting, maintaining a fair balance for the borrower.

Strategic Alignment: These metrics confirm that the HistGradientBoosting approach effectively captures the complex, nonlinear interactions between borrower behavior and regional volatility.

## 8. Explainable Artificial Intelligence Analysis

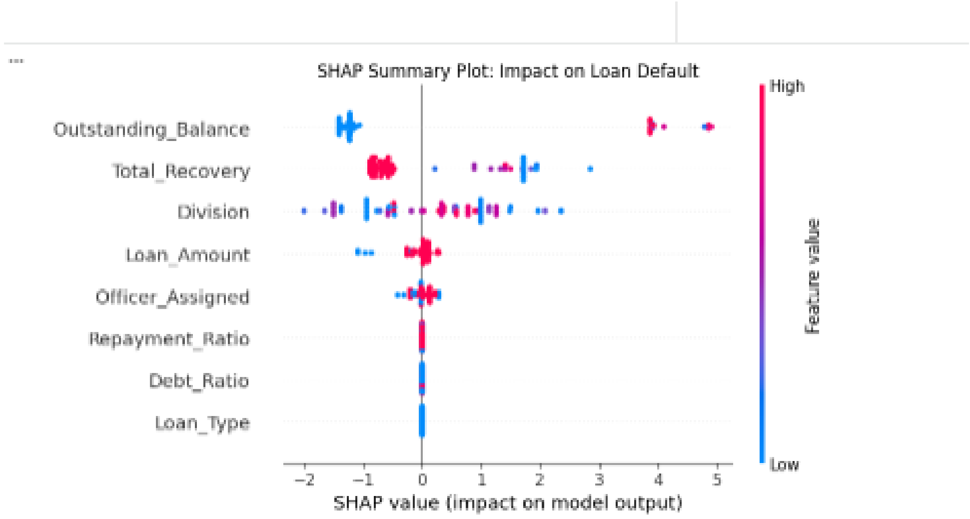


**Figure 10. SHAP summary plot**

Figure 10 presents the SHAP summary plot, which provides a global explanation of the Gradient Boosting model by illustrating the contribution of each feature across all predictions. The horizontal axis represents SHAP values, indicating the magnitude and direction of each feature's impact on default risk. The results show that outstanding balance has the strongest influence on loan default prediction, followed by total recovery and agricultural division. Higher outstanding balances increase default probability, while higher recovery values reduce risk. The mixed distribution of SHAP values for agricultural division highlights the presence of region-specific risk patterns within the agricultural lending portfolio.

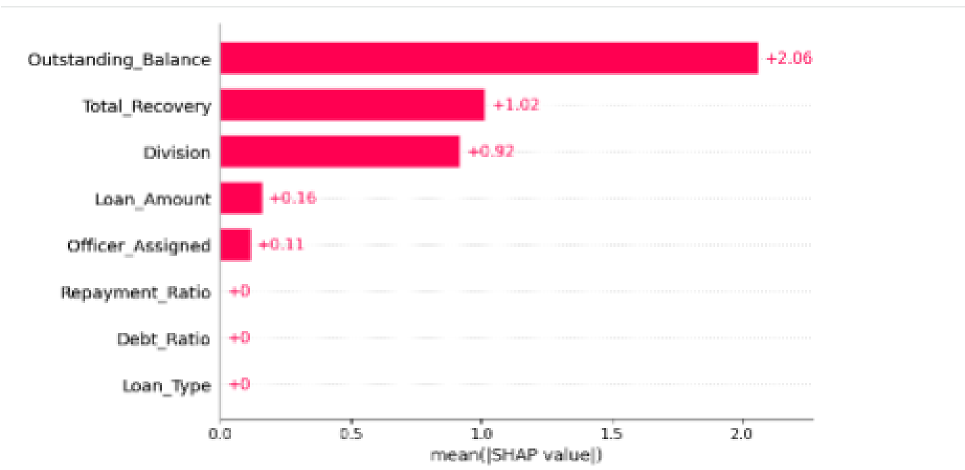


**Figure 11: Mean Absolute SHAP Values**

Figure 11 summarizes the average magnitude of SHAP values for each feature, providing a ranked measure of global feature importance. Outstanding balance emerges as the most influential predictor, indicating that accumulated unpaid obligations are the primary driver of default risk. Total recovery and agricultural division also exhibit substantial contributions, reinforcing the importance of repayment behavior and geographic factors. Features such as loan amount and officer assignment show moderate influence, while repayment ratio, debt ratio, and loan type contribute marginally to the model's predictions.

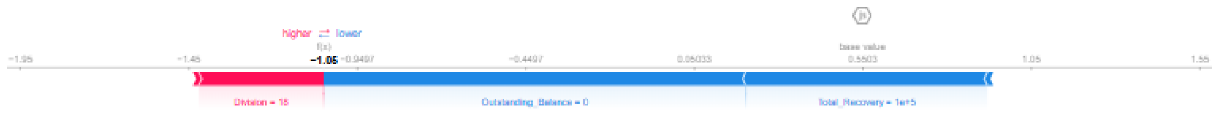


**Figure 12: Local SHAP Explanation**

Figure 12 provides a local explanation of a single loan prediction by decomposing the model output into feature-level contributions. The visualization shows how individual features shift the prediction from the baseline value toward a higher or lower default risk. In this instance, outstanding balance contributes positively to default risk, while total recovery reduces the predicted risk. The agricultural division also plays a notable role, reflecting location-specific factors. This case-level explanation demonstrates the transparency of the proposed model and supports accountable decision-making in loan assessment.

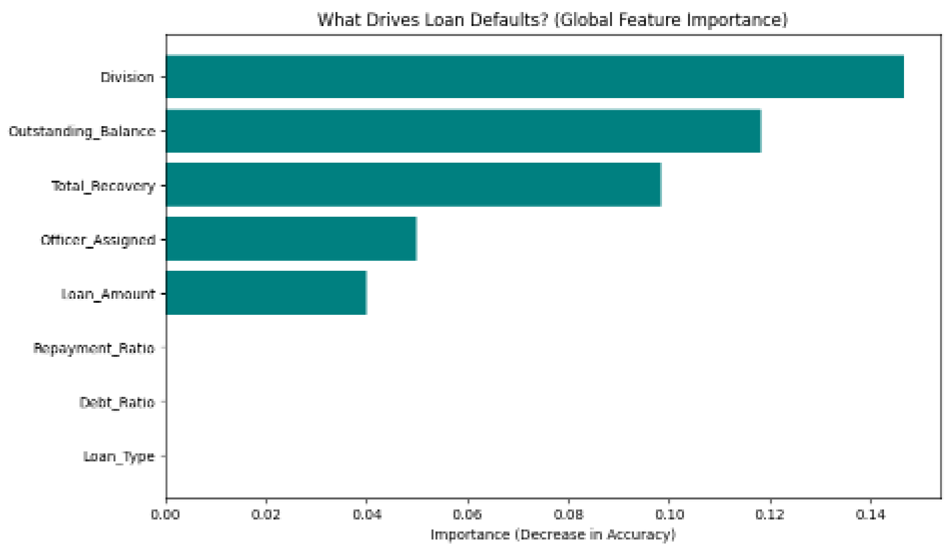


**Figure 13: Permutation-Based Global Feature Importance**

Figure 13 presents permutation-based feature importance, which evaluates the decrease in model accuracy when individual feature values are randomly shuffled. A larger decrease indicates greater importance. The results confirm that outstanding balance, agricultural division, and total recovery are the most critical features influencing model performance. The consistency between permutation importance and SHAP-based explanations strengthens confidence in the robustness and reliability of the identified risk factors.

## 9. Discussion

The results indicate that agricultural loan defaults in Sri Lanka are driven by a combination of financial burden, repayment behavior, geographic conditions, and institutional oversight. The integration of Explainable AI ensures that model predictions are transparent and aligned with real-world financial logic, thereby increasing trust and usability.

## 10. Limitations and Ethical Considerations

This study relies on structured historical data and does not incorporate external factors such as weather conditions or commodity price fluctuations, which are known to influence agricultural income. Additionally, while Explainable AI improves transparency, automated predictions should not replace human judgment. Human oversight remains essential to prevent bias and ensure ethical lending practices, as emphasized in modern financial AI governance frameworks.

## 11. Conclusion and Future Work

This research presents an explainable Gradient Boosting framework for agricultural loan default prediction in Sri Lanka. The proposed model achieves strong predictive performance while maintaining interpretability through Explainable AI. Future work may incorporate climate data, satellite imagery, and real-time monitoring systems to enhance robustness and predictive accuracy.

## References

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. Annals of Statistics, 29(5), 1189–1232.

Kuhn, M., & Johnson, K. (2013). Applied Predictive Modeling. Springer.
Lessmann, S., Baesens, B., Seow, H. V., & Thomas, L. C. (2015). Benchmarking state-of-the-art classification algorithms for credit scoring. European Journal of Operational Research, 247(1), 124–136.
Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30.
Ding, R. (2022). Intelligent financial risk assessment using machine learning. Computational Economics, 59(4), 1335–1354.