Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

Bank Customer Churn Prediction Using CTGAN-Augmented Data and Boosting-Based Ensemble Learning with SHAP Explainable AI

Mohamad Syazimmi Hersyaputra[1], Shintami Chusnul Hidayati*[2]

[1,2]Department of Informatics, Institut Teknologi Sepuluh Nopember, Indonesia

Email: [1]shintami@its.ac.id

Received : Dec 18, 2025; Revised : Jan 5, 2026; Accepted : Jan 19, 2026; Published : Jun 15, 2026

Abstract

Customer churn prediction remains a fundamental concern in the banking domain due to its direct impact on revenue stability and long-term customer value. A key challenge in churn modeling lies in severe class imbalance, which often limits model sensitivity toward minority churn cases. This study aims to develop an integrated and explainable churn prediction framework that effectively addresses class imbalance while maintaining robust predictive performance and interpretability. The proposed approach employs Conditional Tabular Generative Adversarial Networks (CTGAN), comparison of five boosting-based ensemble learning, and SHapley Additive exPlanations (SHAP) to preserve model interpretability. CTGAN is leveraged to synthesize high-fidelity instances for the churn class, yielding a class-balanced dataset that retains intricate tabular feature distributions. Five boosting-based ensemble models, XGBoost, CatBoost, Gradient Boosting Machine (GBM), Stochastic Gradient Boosting (SGB), and LightGBM, are systematically tuned using randomized hyperparameter optimization and evaluated under consistent experimental settings. Model performance is assessed using accuracy, precision, recall, and F1-score to capture classification performance under class imbalance. To ensure transparency, SHAP is applied to analyze global feature importance influencing churn predictions. Experimental results indicate CTGAN enhances model learning stability and detection capability. Among the evaluated models, CatBoost achieves the best results, with an accuracy of 0.9748 and an F1-score of 0.9178. The explainability analysis reveals that transactional features play a dominant role in churn. The novelty of this study lies in a unified and explainable churn prediction framework that integrates CTGAN-data augmentation, boosting ensembles, and interpretability for robust decision support in banking analytics.

Keywords : *Bank Customer Churn, Boosting-Based Ensemble Learning, CTGAN, Explainable AI, SHAP.*

1. INTRODUCTION

The loss of customers through service discontinuation, known as customer churn, has long been recognized as a major concern in the banking industry due to its financial consequences [1]. Alongside this, the evolution of digital banking environments has increasingly influenced how customers interact with banking services [2]. Customers are no longer bound to a single financial service provider [3], as they can easily switch to alternative banks or fintech platforms that offer competitive pricing, seamless digital interfaces, and personalized financial services. This intensified competition has weakened traditional customer loyalty [4], making churn increasingly influenced by service quality, transaction efficiency, mobile banking usability, fee structures, and customer support responsiveness. From a business perspective, unmanaged churn directly affects revenue stability, customer lifetime value, and long-term market positioning [5]. Given that customer acquisition generally requires higher expenditure compared to customer retention [5][6], early and accurate identification of potential churners has become a strategic imperative for banks aiming to design proactive and targeted retention strategies [7].

To address this challenge, Extensive investigations into customer churn prediction have been conducted within the framework of machine learning approaches [8]. Early studies commonly employed Decision Tree models due to their simplicity and interpretability [9], followed by more advanced

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

approaches such as Random Forest (RF) [10], Support Vector Machines (SVM) [11], and Deep Neural Networks [12] to capture more complex customer behavior patterns. Despite these advancements, several critical limitations remain. A major challenge is severe class imbalance, where churned customers represent a small minority compared to active customers [13], thereby favoring the majority class during learning and diminishing sensitivity to churn instances that are most valuable for decision making [14].

More recent works demonstrated the effectiveness of ensemble-based and boosting-based models for churn prediction using tabular banking data, including RF, Logistic Regression, SVM, and XGBoost, achieving reasonable to high predictive accuracy [15][16]. However, these studies typically evaluate models independently, without systematic comparison across multiple boosting-based methods under identical experimental settings, and often prioritize performance metrics over interpretability. In addition, ensemble and voting-based classifiers have been proposed to improve accuracy, yet these approaches largely treat models as black boxes and provide limited insight into feature-level decision mechanisms [17][18]. From a data perspective, most existing studies rely on conventional oversampling techniques such as SMOTE to address class imbalance [16][19], while effective in increasing minority class samples, SMOTE relies on linear interpolation, which may generate noisy samples in majority-class regions [20] and exhibits limited capacity to disentangle intricate interactions across heterogeneous numerical and categorical attributes. Furthermore, the limited adoption of Explainable AI (XAI) techniques in prior works constrains their applicability in real-world banking environments, where transparency, auditability, and regulatory compliance are essential requirements [21].

To overcome these challenges, this study introduces Conditional Tabular Generative Adversarial Networks (CTGAN) as a synthetic data generation mechanism. Unlike traditional oversampling techniques, CTGAN is specifically designed for tabular data and is capable of generating high-fidelity synthetic samples that maintain complex feature dependencies and realistic category distributions [22]. By augmenting the minority churn class using CTGAN, this research creates a more balanced and informative training dataset, enabling boosting-based models to learn churn-related patterns more effectively. The novelty of this work lies in systematically evaluating and comparing XGBoost, CatBoost, Gradient Boosting Machine (GBM), Stochastic Gradient Boosting (SGB), and LightGBM under identical CTGAN-augmented data conditions, allowing for a fair and rigorous assessment of their respective strengths and limitations in churn prediction.

Beyond predictive accuracy, model explainability is a critical requirement for operational deployment in banking environments [21]. Highly accurate black-box models are often unsuitable in practice due to their limited transparency [21]. To address this issue, this study integrates Explainable Artificial Intelligence (XAI) through Shapley Additive exPlanations (SHAP), which provides a theoretically grounded framework for quantifying feature contributions [23]. By integrating SHAP with CTGAN-augmented boosting models, this study identifies the best-performing approach while revealing key factors driving churn predictions, enabling more actionable insights for business analysts.

Overall, this work introduces a comparative framework that addresses three key limitations in existing churn prediction research: (1) handling severe class imbalance using CTGAN-based data augmentation, (2) conducting a structured comparison of five boosting-based ensemble learning models, and (3) embedding explainability through SHAP to ensure transparency and regulatory alignment. By bridging predictive performance, data realism, and interpretability within a unified experimental design, this study contributes a practical and robust churn prediction framework suitable for modern banking environments. Accordingly, this study investigates the following research questions: (RQ1) how does CTGAN-based data augmentation affect the predictive performance of boosting-based ensemble models, (RQ2) which boosting-based model achieves the most balanced performance, and (RQ3) how SHAP-based explainability reveals the key drivers of customer churn in the banking domain.

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

2. METHOD

This research adopts a systematic approach to bank customer churn prediction, beginning with data acquisition and preparation, which includes feature transformation, data scaling, and class balancing. The complete methodological pipeline is presented graphically in Figure 1, while explainable analysis is incorporated at the final stage to provide transparent and interpretable predictive outcomes.

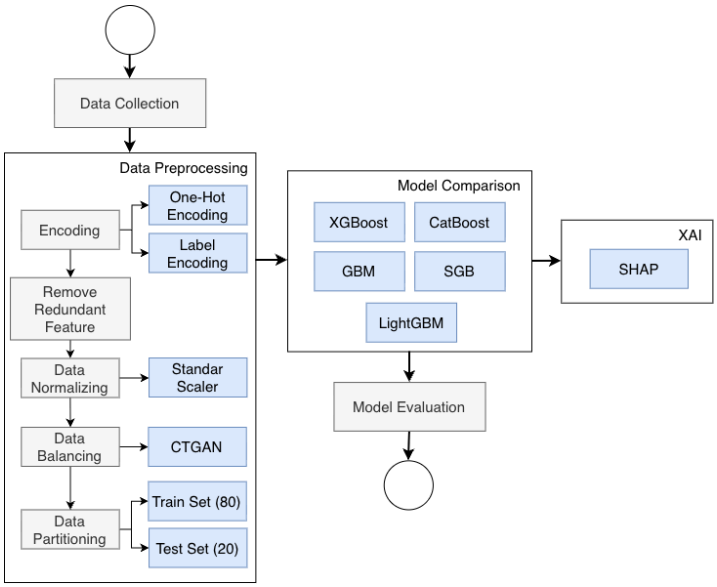


Figure 1. Proposed Research Methodology

2.1. Data Preparation Overview

This study is based on an open-access Credit Card Customers dataset provided via the Kaggle repository. As shown in Figure 2, the target variable exhibits an imbalanced distribution. The dataset consists of 10,127 records and 21 features.

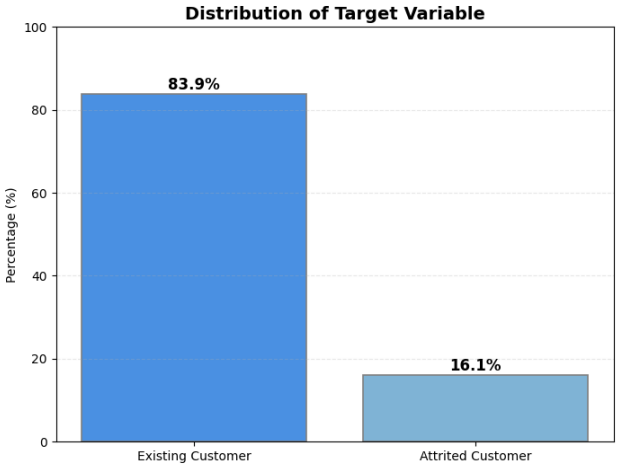


Figure 2. Distribution of Target Variable

The dataset features are categorized into identification, demographic, account relationship and activity, and transaction-related attributes. The identification feature includes *CLIENTNUM* (numerical, unique customer identifier). Demographic features consist of *Customer_Age* (numerical), Gender (categorical), *Dependent_count* (numerical), *Education_Level* (categorical), *Marital_Status*

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

(categorical), *Income_Category* (categorical), and *Card_Category* (categorical). Account relationship and activity features include *Months_on_book* (numerical), *Total_Relationship_Count* (numerical), *Months_Inactive_12_mon* (numerical), and *Contacts_Count_12_mon* (numerical). Transactional and financial features capture spending and credit behavior, comprising *Credit_Limit* (numerical), *Total_Revolving_Bal* (numerical), *Avg_Open_To_Buy* (numerical), *Avg_Utilization_Ratio* (numerical), *Total_Trans_Amt* (numerical), *Total_Trans_Ct* (numerical), *Total_Amt_Chng_Q4_Q1* (numerical), and *Total_Ct_Chng_Q4_Q1* (numerical).

2.2. Data Preprocessing

The data preprocessing phase was performed to prepare the dataset for model training by converting categorical features into numerical form. Two distinct encoding approaches were selected to accommodate differing feature characteristics: One-Hot Encoding for nominal features without inherent order and Label Encoding for features that exhibit ordinal properties or do not require extensive column expansion. In this study, *Gender* and *Marital_Status* were encoded using One-Hot Encoding, while *Education_Level*, *Income_Category*, and *Card_Category* were transformed using Label Encoding. This approach ensures appropriate numerical representation of categorical variables while preserving their semantic meaning.

The feature selection stage focuses on removing redundant variables to reduce multicollinearity and improve model robustness. In this study, *Avg_Open_To_Buy* was excluded due to its strong similarity with *Credit_Limit*, as both represent customers' available credit capacity. As illustrated in Figure 3, the two features show highly similar distribution patterns across customer groups, indicating overlapping information. Consequently, *Avg_Open_To_Buy* was removed to enhance model stability.

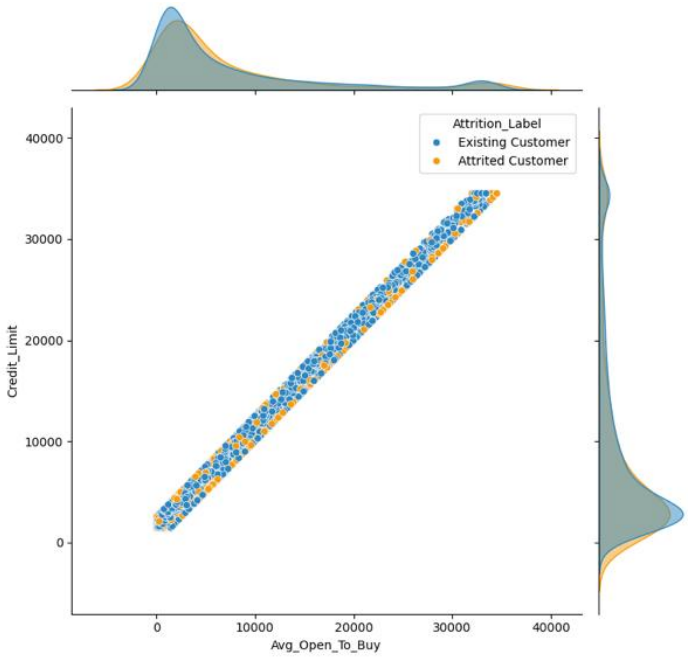


Figure 3. Distribution of *Avg_Open_To_Buy* and *Credit_Limit* Feature against Target Variable

Following feature selection, data normalization was applied to ensure that numerical features were on a comparable scale prior to model training. In this study, *StandardScaler* was applied to normalize numerical features by rescaling them to zero mean and unit variance. This normalization process helps prevent features with larger numerical ranges from dominating the learning process and improves the stability of the optimization procedure, particularly for boosting-based models. By

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

standardizing the feature space, the models are able to learn more effectively from balanced feature contributions, leading to improved convergence and more reliable predictive performance.

Subsequently, a data balancing phase is conducted to handle the skewed class distribution inherent in the bank customer churn dataset, where the proportion of customers who have terminated the service is considerably lower than that of retained customers. This condition may bias the learning process toward the majority class, reducing sensitivity to minority instances. To address this issue, this study applies CTGAN, a GAN-based method specifically developed to synthesize realistic tabular data while maintaining the underlying feature distributions. CTGAN comprises two core components, namely a generator and a discriminator [24]. The generator creates artificial samples from random noise z conditioned on categorical variables c, whereas the discriminator attempts to differentiate between real and generated data samples [25]. The learning process is formulated as a min–max optimization problem, as defined by the objective function presented in Equation 1 [26].

$$\min_G \max_D \mathbb{E}_{X\sim P_{data}}[\log D(x)] + \mathbb{E}_{z\sim P_z, c\sim P_c}[\log 1 - D(G(z,c)] \quad (1)$$

Where x represents real data samples, z is Gaussian noise, and c is randomly selected categorical conditions. Through this conditional mechanism, CTGAN is able to generate synthetic samples that specifically represent the minority churn class, resulting in a more balanced target distribution.

For continuous features, CTGAN applies mode-specific normalization to effectively handle multimodal distributions. Each continuous value $C_{i,j}$ is represented with respect to its k-th mode using the following formulation in Equation 2 [26].

$$\alpha_{i,j} = \frac{c_{i,j} - \eta_k}{4\Phi_k} \quad (2)$$

Here η_k represents the average value, while Φ_k denote standard deviation of the k-th mode, respectively. This approach enables CTGAN to model continuous features with complex, multimodal characteristics more accurately.

For discrete features, CTGAN constructs a conditional vector based on one-hot encoding, which is provided as an explicit condition to the generator. The conditional vector is defined as shown in Equation 3.

$$m = m_i^k \oplus \cdots \oplus m_i^{|D_i|} \quad (3)$$

Where m_i^k equals 1 for the selected categorical condition and 0 for the remaining categories. The generator is then guided to produce samples that match the specified condition through a cross-entropy–based penalty function. By leveraging this conditional generation strategy, CTGAN can generate rare-category samples in a controlled manner, effectively balancing the churn class without distorting the natural structure of the dataset. The CTGAN-augmented dataset yields a more balanced and informative training distribution, allowing boosting-based models to capture churn characteristics more effectively, alleviate dominance of the majority class, and enhance responsiveness to churn-indicative patterns.

The preprocessing phase concludes by partitioning the dataset into training and testing (8:2) subsets. This split enables unbiased evaluation using unseen data, in which model fitting is conducted on the training data and evaluation is performed on the testing data. The chosen ratio balances learning adequacy with reliable evaluation of the model's generalization capability.

2.3. Boosting-Based Ensemble Modeling

Multiple boosting-based ensemble learners are comparatively assessed in this study for bank customer churn prediction. Boosting refers to a staged ensemble strategy in which successive models

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

are trained to emphasize previously mispredicted instances, enabling gradual improvement in overall accuracy. Boosting-based methods are particularly suitable for tabular financial data due to their ability to model non-linear relationships, feature interactions, and heterogeneous feature distributions. In this research, five representative boosting models are evaluated, namely XGBoost, CatBoost, GBM, SGB, and LightGBM, each reflecting different design philosophies within the boosting framework.

The selected models differ in their learning mechanisms and optimization strategies. XGBoost represents a regularized gradient boosting approach that incorporates first- and second-order gradient information along with explicit regularization to control model complexity [27]. GBM follows the classical gradient boosting paradigm by iteratively fitting decision trees to residual errors [28], while SGB extends GBM by introducing stochastic subsampling of training instances, which helps reduce variance and improve generalization [29]. LightGBM adopts a leaf-wise tree growth strategy combined with histogram-based discretization, enabling efficient learning on large-scale tabular data [30]. CatBoost, on the other hand, is designed to handle categorical features more robustly through ordered target statistics and permutation-driven learning, reducing target leakage and improving stability [28]. The main conceptual characteristics differentiating the boosting models used in this study are presented in Table 1.

Table 1. Taxonomy of Boosting-Based Ensemble Models Evaluated in This Study

Model	Boosting Type	Core Idea
XGBoost	Regularized Gradient Boosting	Regularization with gradient-based learning
CatBoost	Categorical Gradient Boosting	Native categorical feature handling
GBM	Gradient Boosting	Residual learning via gradients
SGB	Stochastic Gradient Boosting	Gradient Boosting with subsampling
LightGBM	Leaf-Wise Gradient Boosting	Leaf-wise tree growth strategy

Table 2. Hyperparameter Search Spaces

Model	Tuned Hyperparameters (Search Space)
XGBoost	*n_estimators* ∈ {50, 100, 200}; *max_depth* ∈ {3, 5, 7}; *learning_rate* ∈ {0.01, 0.1, 0.2}; *subsample* ∈ {0.7, 0.8, 1.0}; *colsample_bytree* ∈ {0.7, 0.8, 1.0}; *gamma* ∈ {0, 0.1, 0.2}
CatBoost	*iterations* ∈ {200, 500, 800}; *depth* ∈ {3, 5, 7}; *learning_rate* ∈ {0.01, 0.1, 0.2}; *l2_leaf_reg* ∈ {1, 3, 5, 7, 9}; *random_strength* ∈ {1, 3, 5}; *bagging_temperature* ∈ {0, 0.5, 1.0}
GBM	*n_estimators* ∈ {100, 200, 300}; *learning_rate* ∈ {0.01, 0.1, 0.2}; *max_depth* ∈ {3, 5, 7}; *subsample* ∈ {0.7, 0.8, 1.0}; *min_samples_split* ∈ {2, 5, 10}; *min_samples_leaf* ∈ {1, 3, 5}
SGB	*n_estimators* ∈ {100, 200, 300}; *learning_rate* ∈ {0.01, 0.1, 0.2}; *max_depth* ∈ {3, 5, 7}; *subsample* ∈ {0.7, 0.8, 0.9}; *min_samples_split* ∈ {2, 5, 10}; *min_samples_leaf* ∈ {1, 3, 5}
LightGBM	*n_estimators* ∈ {100, 200, 400, 600}; *learning_rate* ∈ {0.01, 0.05, 0.1, 0.2}; *max_depth* ∈ {−1, 3, 5, 7}; *num_leaves* ∈ {15, 31, 63, 127}; *subsample* ∈ {0.7, 0.8, 0.9, 1.0}; *colsample_bytree* ∈ {0.7, 0.8, 0.9, 1.0}; *min_child_samples* ∈ {10, 20, 30, 50}; *reg_alpha* ∈ {0.0, 0.1, 0.5, 1.0}; *reg_lambda* ∈ {0.0, 0.1, 0.5, 1.0}

The selected parameter ranges are designed to reflect the core learning mechanisms of each model while maintaining a comparable level of search complexity across experiments. By explicitly defining consistent and representative search spaces, this study ensures that each model is optimized under fair and controlled conditions prior to performance evaluation. Table 2 summarizes the hyperparameter search spaces defined for each boosting-based ensemble model evaluated in this study.

By applying consistent tuning strategies and comparable parameter search spaces, this experimental design ensures that performance differences among models primarily reflect their intrinsic learning mechanisms rather than suboptimal hyperparameter configurations. This approach enables a systematic assessment of how different boosting paradigms respond to CTGAN-augmented data in the context of bank customer churn prediction.

To ensure a fair and robust comparison, hyperparameter optimization is conducted for each boosting model using *RandomizedSearchCV*. This approach is chosen due to its computational efficiency and effectiveness in exploring large hyperparameter spaces. The search process evaluates randomly sampled parameter combinations over 20 iterations using 5-fold cross-validation, with the F1-score as the optimization metric to account for class imbalance in the churn dataset.

2.4. Model Evaluation

Model performance was quantitatively assessed through four classification metrics, namely accuracy, precision, recall, and F1-score, which together offer a thorough performance assessment, particularly in the presence of skewed class distributions. All evaluation metrics were derived from the confusion matrix, which comprises True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) outcomes. Accuracy reflects the overall correctness of model predictions and is formally expressed in Equation [4].

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Precision measures the reliability of churn predictions by calculating the fraction of correctly identified churn cases relative to all instances predicted as churn, as defined in 5 [30].

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Recall (sensitivity), evaluates the model's effectiveness in detecting actual churn, a metric of particular importance for the minority class in imbalanced datasets, as formulated in Equation 6 [30].

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

Finally, the F1-score combines precision and recall into a single balanced metric, making it suitable for evaluating performance under class imbalance, as shown in Equation 7 [30].

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (7)$$

2.5. Explainable AI (XAI)

To guarantee that the churn prediction model achieves both reliable prediction accuracy and explainable, auditable decision-making, this study incorporates XAI through SHAP. SHAP is an interpretation method derived from Shapley value principles in cooperative game theory, which quantifies the individual contribution of each input feature to the model's predictions [31]. In the SHAP framework, the original complex prediction function $f(x)$ is represented by an additive surrogate explanation model, as formalized in Equation 8 [31].

$$f(x) = g(z) = \phi_0 + \sum_{i=1}^{M} \phi_i z_i \quad (8)$$

Here $z = \{0,1\}^M$ represents a binary vector indicating the presence of features, ϕ_0 denotes the base value corresponding to the expected model output, and ϕ_i represents the contribution of the i-th feature. This additive formulation enables consistent and measurable feature attribution at both global

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

and local levels, providing transparent explanations that support risk analysis and regulatory compliance in banking applications.

3. RESULT

This section reports the empirical findings obtained from the proposed bank customer churn prediction framework. The effectiveness of the boosting-based ensemble approaches is systematically analyzed and benchmarked using several classification indicators, enabling a fair and consistent comparison of their predictive capability. Furthermore, the reported results offer analytical insights into how data balancing strategies and model choice influence churn prediction outcomes.

3.1. Result of Feature Encoding

This subsection reports the outcomes of the feature encoding stage conducted on categorical attributes before the model learning process. Feature encoding is applied to convert categorical data into numeric formats that are suitable for processing by boosting-based learning algorithms. As an illustration, Table 3 presents the One-Hot Encoding results for the *Gender* variable. Prior to encoding, this attribute contains two nominal classes, namely Male and Female. After transformation, where categorical membership is represented by a binary indicator, with 1 corresponding to inclusion and 0 to exclusion. In this representation, Male is encoded as (1, 0), while Female is encoded as (0, 1). This transformation maintains the categorical semantics of the original feature and avoids imposing any unintended ordinal structure, thereby preserving the interpretability and integrity of the *Gender* attribute for subsequent modeling phases.

Table 3. One-Hot Encoding Results for the *Gender* Feature

Before Encoding	After Encoding	
	Male	Female
Male	1	0
Female	0	1

Table 4 shows the results of Label Encoding applied to the *Card_Category* feature as an example. Each categorical value is mapped to a unique numerical label, allowing the feature to be represented in a compact numeric form suitable for model training. This encoding approach avoids increasing feature dimensionality and is well-suited for tree-based boosting models, which can effectively handle discrete numeric values without assuming an ordinal relationship.

Table 4. Label Encoding Results for the *Card_Category* Feature

Before Encoding	After Encoding
Blue	0
Silver	1
Gold	2
Platinum	3

3.2. Results of Data Normalization Using *StandardScaler*

This subsection presents the results of the data normalization process applied prior to model training. A normalization procedure is applied to align numerical features onto a consistent scale, reducing the risk of large-magnitude variables exerting undue influence on boosting-based learning algorithms. In this study, normalization is conducted using the *StandardScaler* method, by re-centering each numerical attribute to a zero-mean distribution and scaling it to unit variance.

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

Figure 4 illustrates the two-dimensional PCA visualization of data distributions before and after normalization. Prior to normalization, the data points exhibit a highly skewed and dispersed distribution, particularly along the first principal component, indicating the dominance of features with large magnitudes. After applying *StandardScaler*, the data distribution becomes more compact and evenly spread across both principal components, reflecting improved feature scaling. This transformation improves numerical robustness and supports more efficient model optimization by placing all variables on a comparable scale, which is particularly crucial for gradient-driven boosting methods.

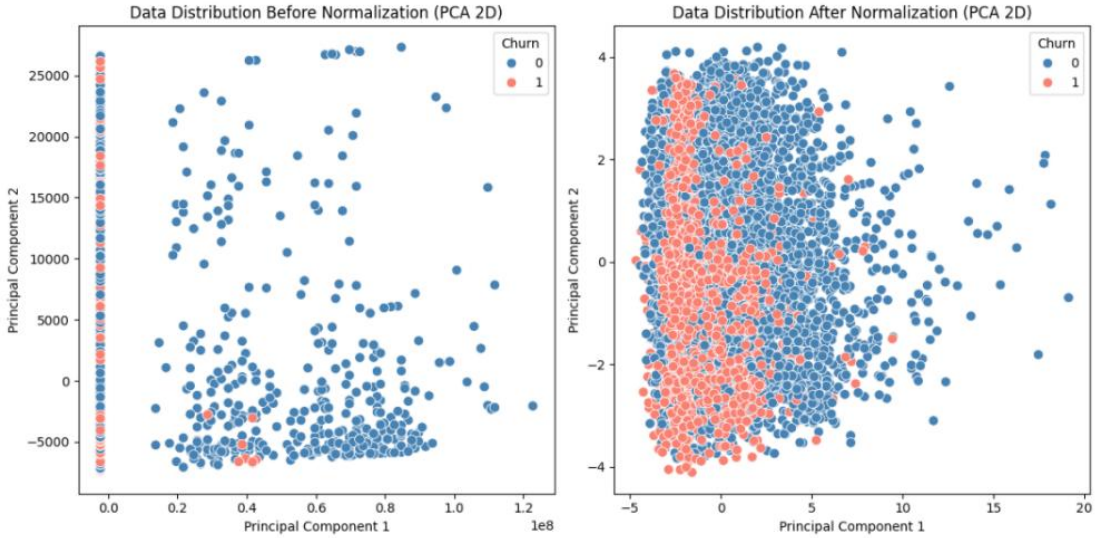


Figure 4. Data Distribution Before and After Data Normalization

3.3. Results of CTGAN-Based Data Balancing

This subsection reports the results of the data balancing process conducted using CTGAN. Data rebalancing is employed to mitigate the skewed class distribution in the churn dataset, in which non-churn instances dominate churn cases and may skew the learning process toward the majority class during training.

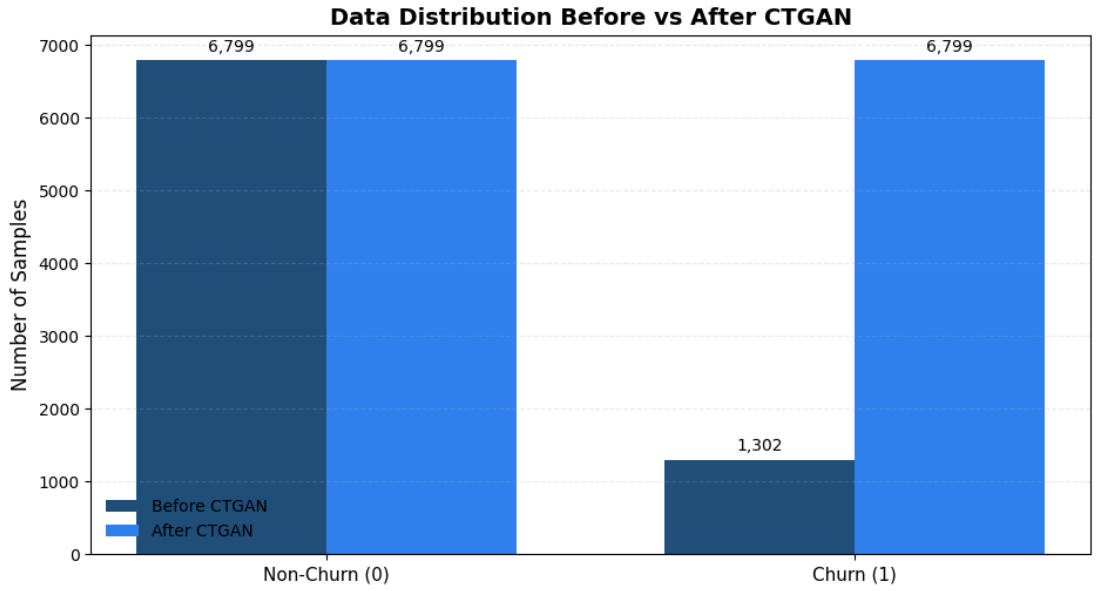


Figure 5. Data Distribution of CTGAN-Augmented Data

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

Figure 5 highlights the uneven class proportions present prior to the CTGAN augmentation process, with 6,799 non-churn samples compared to only 1,302 churn samples. After CTGAN-based augmentation, the number of churn samples is increased to match the non-churn class, resulting in a balanced distribution of 6,799 samples for each class. Beyond achieving numerical balance, CTGAN generates distribution-aware synthetic samples that preserve complex relationships among numerical and categorical features, enabling the models to learn representative churn patterns rather than relying on simple interpolation. This balanced dataset provides a more equitable learning environment for the boosting-based models by ensuring sufficient representation of the minority class. Consequently, the application of CTGAN helps mitigate class bias and supports more reliable learning of churn-related patterns in subsequent modeling stages.

3.4. Results of Hyperparameter Tuning

This subsection outlines the hyperparameter optimization results obtained for boosting-based ensemble models via *RandomizedSearchCV* with 5-fold cross-validation, optimized toward the F1-score. Across all boosting-based ensemble configurations except LightGBM, 20 parameter configurations were evaluated, resulting in a total of 100 model fits per algorithm. The tuning process aims to identify optimal parameter combinations that balance model complexity and generalization performance on the churn prediction.

The optimal hyperparameter configurations obtained from the tuning process for each boosting-based model are summarized in Table 5. The selected parameters reflect the distinct learning mechanisms and regularization strategies of each model. XGBoost and LightGBM benefit from a combination of subsampling and explicit regularization to control model complexity, while CatBoost achieves stable performance through depth control and strong L2 regularization without additional bagging. In contrast, GBM and SGB rely on tree depth and leaf constraints, with SGB further incorporating stochastic subsampling to enhance generalization. Overall, the tuned configurations ensure that each model is evaluated under its most suitable settings, enabling a fair and reliable performance comparison in the subsequent analysis.

Table 5. Optimal Hyperparameter Configurations for Boosting-Based Models

Model	Optimal Hyperparameters
XGBoost	*n_estimators* = 200; *max_depth* = 7; *learning_rate* = 0.1; *subsample* = 0.8; *colsample_bytree* = 0.7; *gamma* = 0.1
CatBoost	*iterations* = 500; *depth* = 5; *learning_rate* = 0.1; *l2_leaf_reg* = 9; *random_strength* = 3; *bagging_temperature* = 0
GBM	*n_estimators* = 200; *max_depth* = 7; *learning_rate* = 0.1; *subsample* = 1.0; *min_samples_leaf* = 5
SGB	*n_estimators* = 200; *max_depth* = 5; *learning_rate* = 0.1; *subsample* = 0.8; *min_samples_leaf* = 3
LightGBM	*n_estimators* = 400; *num_leaves* = 63; *max_depth* = 5; *learning_rate* = 0.1; *subsample* = 0.7; *colsample_bytree* = 0.8; *reg_alpha* = 1.0; *reg_lambda* = 1.0

3.5. Model Performance Comparison

This subsection examines the predictive capabilities of the proposed boosting-based ensemble approaches by employing four common classification measures. These metrics are jointly examined to provide a more complete interpretation of model performance, especially under imbalanced churn conditions where dependence on a single metric may produce biased interpretations. Table 6 consolidates the comparative performance results across the evaluated models.

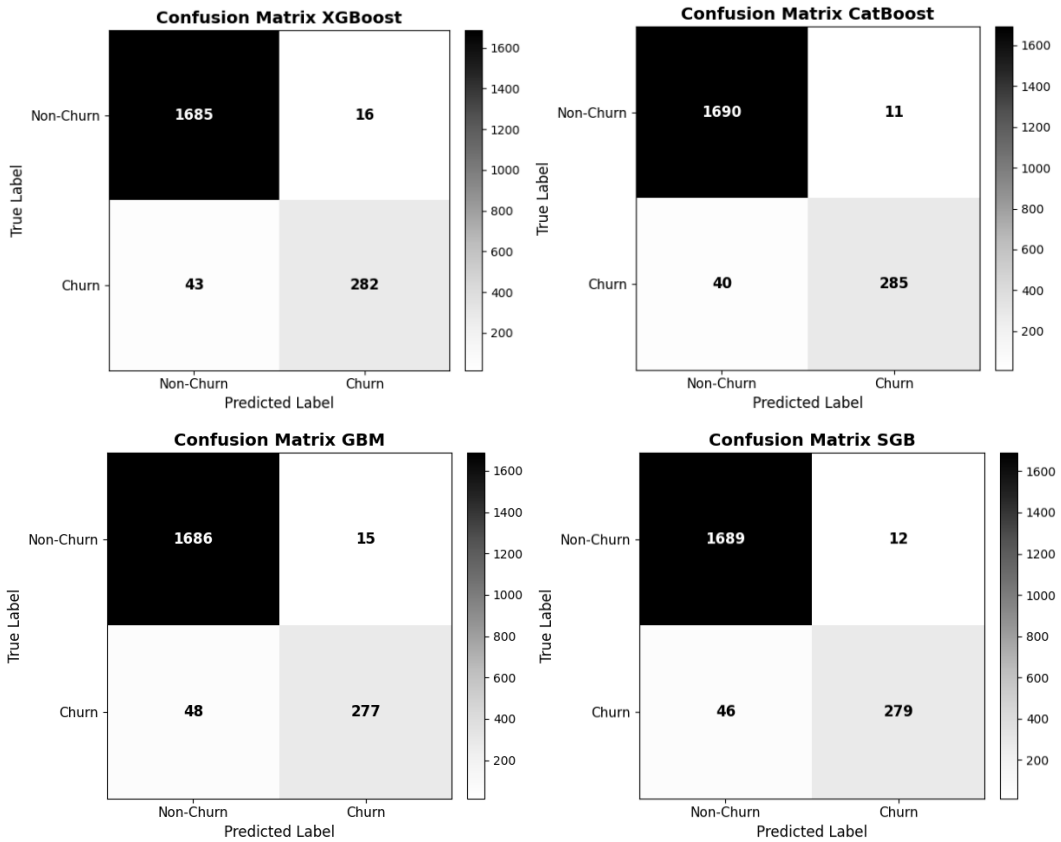
Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

Table 6. Performance Metric Comparison

Model	Accuracy	Precision	Recall	F1-Score
XGBoost	0.9708	0.9463	0.8676	0.9052
CatBoost	**0.9748**	**0.9628**	**0.8769**	**0.9178**
GBM	0.9689	0.9486	0.8523	0.8978
SGB	0.9713	0.9587	0.8584	0.9058
LightGBM	0.9708	0.9493	0.8646	0.9049

The findings show that CatBoost provides the most balanced and consistently strong performance. This can be linked to its ordered boosting and robust categorical handling, which mitigate prediction shift and overfitting while capturing feature interactions typical in tabular banking data. XGBoost and LightGBM achieve similar overall accuracy and F1-scores, but differ in operating point: XGBoost yields slightly higher recall, consistent with its regularized gradient boosting that tends to recover more churn cases, whereas LightGBM attains slightly higher precision with lower recall, aligning with its leaf-wise growth that can form more selective decision boundaries. SGB attains the highest precision with a competitive F1-score, suggesting that stochastic subsampling improves generalization and reduces false positives, though at a small recall cost. In contrast, GBM has the lowest recall and F1-score, indicating more missed churn cases, which may reflect lower adaptability to class imbalance under the chosen depth/leaf constraints. Overall, the results highlight how differences in boosting strategy, tree growth, and regularization drive distinct precision–recall trade-offs, making model choice dependent on whether the priority is minimizing false alarms or capturing as many churn cases as possible.

Figure 6 presents the confusion matrices of all evaluated boosting models, illustrating their classification outcomes in distinguishing churn and non-churn customers.



Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

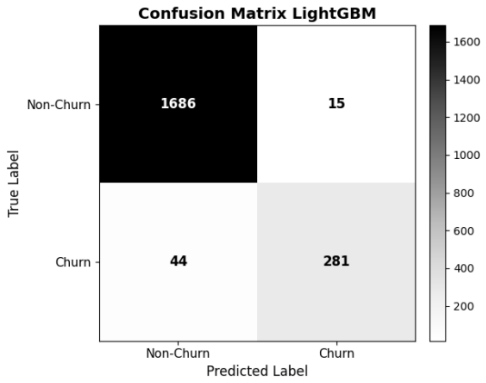


Figure 6. Confusion Matrix Results for Boosting-Based Churn Prediction Models

The confusion matrices illustrate distinct classification behaviors across the evaluated boosting models. CatBoost achieves the highest true positive count with relatively few false positives, resulting in superior recall and F1-score, which can be attributed to its ordered boosting and robust regularization that enhance learning on minority churn samples. XGBoost and LightGBM display similar error patterns, with slightly higher false negatives, underscoring the compensatory interplay between precision and recall arising from their regularized gradient boosting and leaf-wise tree growth strategies. SGB lowers false positives through stochastic subsampling, but sacrifices recall by missing more churn cases, while GBM produces the highest false negatives, indicating a weaker capability in detecting churn customers. These results demonstrate that differences in boosting mechanisms lead to distinct error distributions, emphasizing the need to align model choice with churn detection objectives.

3.6. Result of SHAP Global Feature Importance

This subsection presents a global interpretability analysis of the proposed banking domain churn prediction model using SHAP to identify the most influential features contributing to model decisions. The analysis aims to provide transparency into how key customer attributes affect churn predictions, supporting model interpretability in a banking context.

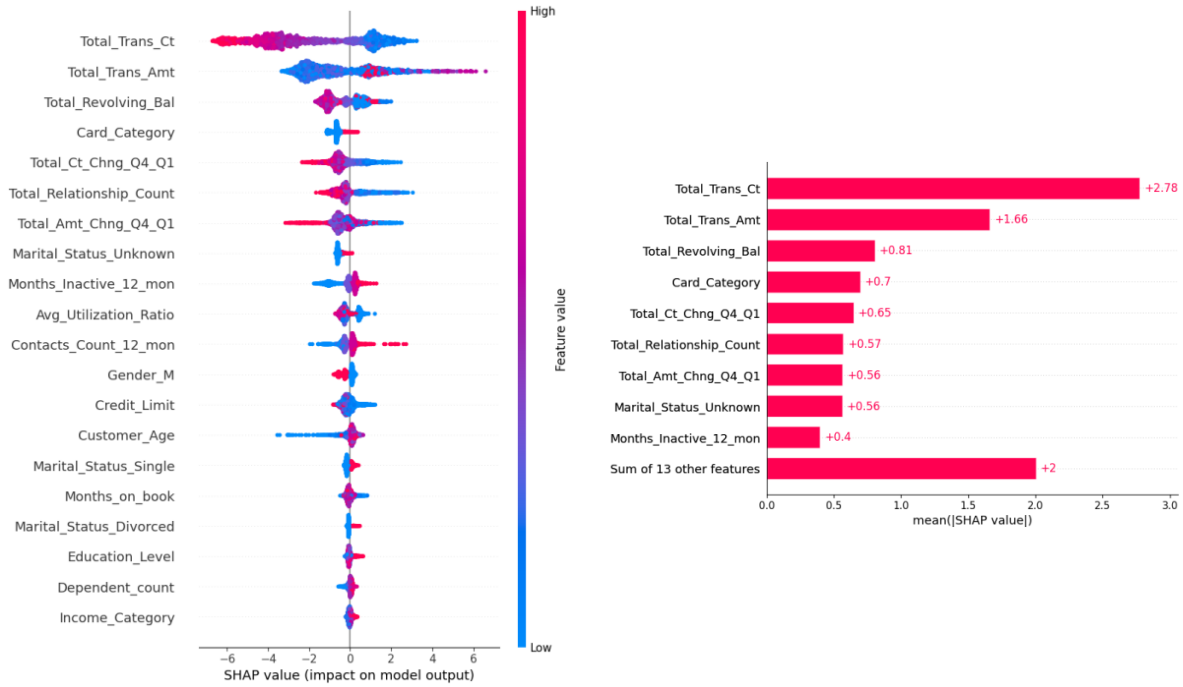


Figure 7. SHAP Feature Importance of XGBoost

Jurnal Teknik Informatika (JUTIF)
P-ISSN: 2723-3863
E-ISSN: 2723-3871
Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

First, The global SHAP key contrubution is first depicted in **Kesalahan! Sumber referensi tidak ditemukan.** for the XGBoost model, indicating that transaction-related variables, particularly *Total_Trans_Ct* and *Total_Trans_Amt*, have the strongest influence on churn predictions. Features related to credit usage and customer engagement, such as *Total_Revolving_Bal*, *Total_Relationship_Count*, and *Months_Inactive_12_mon*, also contribute significantly, while demographic attributes have relatively lower impact. Overall, the results confirm that churn decisions are primarily driven by behavioral and transactional patterns rather than static customer characteristics.

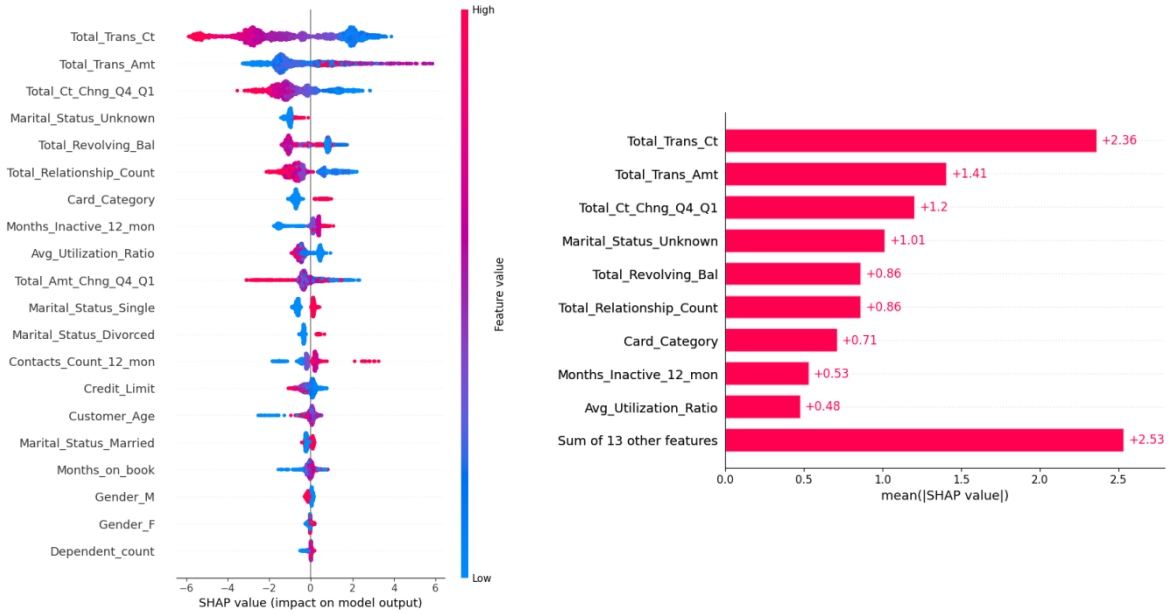


Figure 8. SHAP Feature Importance of CatBoost

Second, Figure 8 indicates that CatBoost's churn predictions are primarily driven by transaction activity (*Total_Trans_Ct* and *Total_Trans_Amt*), followed by changes in transaction behavior and account engagement features, while demographic variables contribute relatively less to the overall model decision.

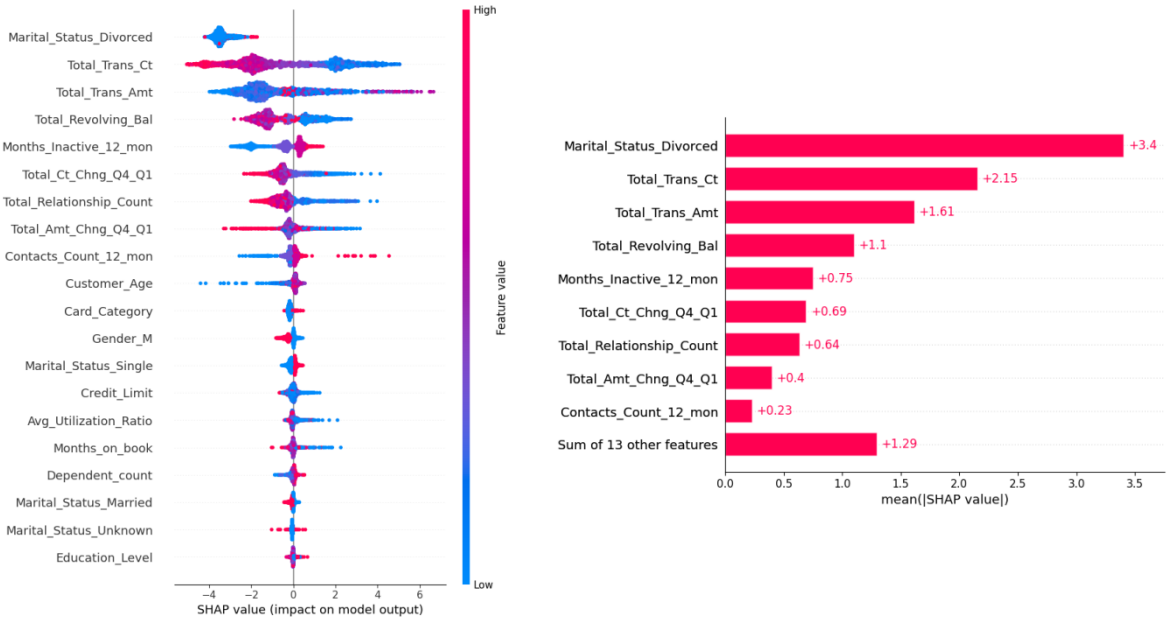


Figure 9. SHAP Feature Importance of GBM

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

Third, Figure 9 shows that the GBM model is strongly influenced by behavioral and account status features, with *Marital_Status_Divorced* emerging as the most dominant contributor, followed by *Total_Trans_Ct* and *Total_Trans_Amt*. Transaction intensity, revolving balance, and inactivity duration further shape the model's predictions, while demographic attributes contribute marginally. This pattern indicates that GBM relies on a combination of transactional behavior and selected categorical indicators to distinguish churn risk, with less emphasis on purely demographic factors.

Fourth, consistent with the GBM results, Figure 10 shows that the SGB model is mainly influenced by *Marital_Status_Divorced* and *Total_Trans_Ct*. The stochastic subsampling mechanism slightly redistributes feature contributions, reducing the dominance of individual features while preserving a similar overall importance pattern.

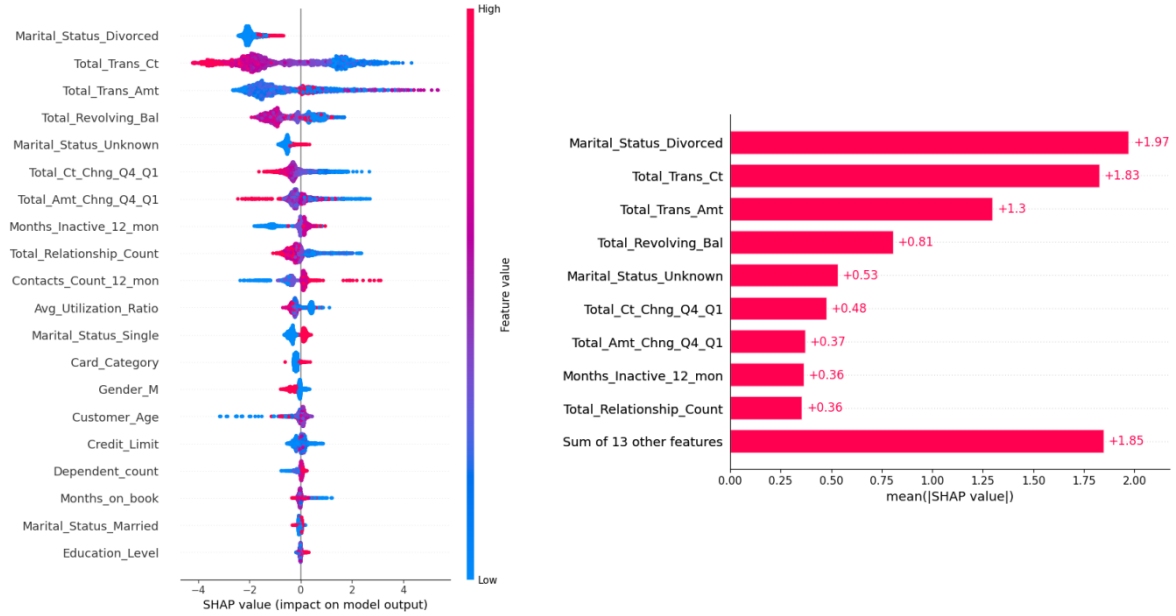


Figure 10. SHAP Feature Importance of SGB

Lastly, Figure 11 shows that LightGBM is mainly driven by *Total_Trans_Ct* and *Total_Trans_Amt*, indicating a strong reliance on transaction-related behavior, consistent with its leaf-wise tree growth strategy.

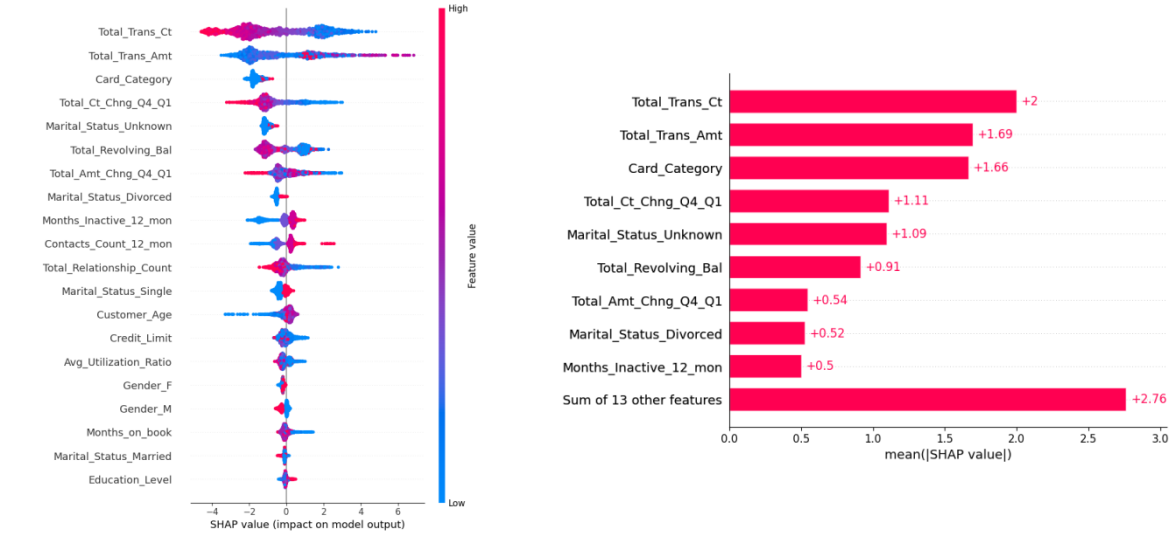


Figure 11. SHAP Feature Importance of LightGBM

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

4. DISCUSSIONS

This study demonstrates that boosting-based ensemble models, when combined with systematic hyperparameter tuning and appropriate CTGAN data balancing, can deliver strong and stable performance. The results of the conducted experiments demonstrate that CatBoost delivers superior predictive performance, with an accuracy of 0.9748 on the Credit Card Customer dataset, indicating its effectiveness in capturing complex, non-linear relationships while remaining robust to class imbalance through CTGAN-based data augmentation. The unified evaluation framework used in this study allows a fair comparison across multiple boosting variants, highlighting how differences in boosting strategy and regularization influence predictive behavior.

Table 7. Performance Comparison of Bank Churn Prediction in Related Works

Method	Dataset Used	Accuracy
Random Forest [15]	Credit Card Customer	0.8870
Random Forest + SMOTE [16]	Credit Card Customer	0.9600
XGBoost [17]	Credit Card Customer	0.9669
Voting Ensemble + SMOTE [18]	Bank Customer Churn Prediction	0.9000
Decision Tress + Feature Selection [19]	Neobank Dataset	0.8050
CatBoost + CTGAN (Best Model in This Study)	Credit Card Customer	**0.9748**

A comparative analysis with prior studies, as summarized in Table 7, highlights several key distinctions that emphasize the novelty of the proposed framework. Earlier churn prediction approaches based on traditional ensemble methods such as RF generally report accuracy values below 0.90, highlighting its limited proficiency in modeling complex non-linear dependencies in banking churn dataset. Although SMOTE-based oversampling has been shown to improve predictive performance, its reliance on linear interpolation limits its ability to preserve higher-order feature dependencies and conditional distributions, which are common in mixed-type banking datasets. In contrast, the proposed framework leverages CTGAN, represents a methodological enhancement, to generate distribution-aware synthetic churn samples, enabling the models to learn richer structural and conditional patterns. Compared to previous XGBoost-based studies reporting accuracies close to 0.97, this study demonstrates a modest but consistent performance gain that arises from the cohesive integration of the proposed framework rather than from model choice alone. Among the evaluated classifiers, the CatBoost–CTGAN combination consistently achieves the most advantageous performance result, attaining the most compelling accuracy–F1 performance gains, which can be attributed to the complementary interaction between CTGAN-based balancing and CatBoost's ordered boosting strategy with native categorical feature handling, reducing target leakage, stabilizing learning, and mitigating prediction bias. Collectively, these results demonstrate that meaningful performance improvements in churn prediction arise from the coordinated integration of data-level augmentation and model-level optimization, positioning the proposed approach as a methodologically consistent advancement over existing studies and positioning the proposed CatBoost–CTGAN approach as a methodologically robust progression beyond existing studies.

Importantly, the evidence suggests that the value of this work goes beyond achieving higher accuracy, but in providing a more comprehensive and methodologically consistent framework. By systematically comparing multiple boosting variants under the same experimental setting, this study offers clearer insights into how different boosting strategies behave on imbalanced banking datasets. More importantly, the integration of SHAP constitutes a key methodological contribution by enabling model-agnostic, fine-grained interpretability across all evaluated boosting models. Rather than treating explainability as a post-hoc visualization, SHAP is systematically employed to reveal consistent global

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

feature importance patterns and to validate that model decisions are primarily driven by transactional and behavioral attributes that align with domain knowledge.

This strengthens the reliability and auditability of the proposed framework, which is essential for regulatory compliance and decision support in banking applications. Furthermore, the integration of SHAP enhances interpretability, enabling the results to be more transparent and actionable in a real banking context. Taken together, The methodological contribution of this study is reflected in a unified framework that integrates CTGAN-based distribution-aware data augmentation, systematic comparison of boosting-based ensemble models, and SHAP-driven explainability. This integration provides a robust and interpretable churn prediction approach that is particularly relevant for decision support and regulatory-compliant analytics in the banking domain.

5. CONCLUSION

This study proposes an integrated framework for bank customer churn prediction that combines CTGAN-based data augmentation, boosting-based ensemble learning comparison, and SHAP-based XAI. Experimental results show that CTGAN effectively enhances data balancing by generating realistic, distribution-preserving synthetic churn samples, leading to improved learning stability for minority classes across all evaluated models. Under a unified experimental setting with consistent preprocessing and hyperparameter tuning, CatBoost yields the strongest observed performance, with accuracy and F1-score values of 0.9748 and 0.9178, surpassing the performance of XGBoost, GBM, SGB, and LightGBM. This superior performance is largely attributable to CatBoost's ordered boosting scheme and effective regularization, which mitigate overfitting and improve generalization when learning from imbalanced tabular banking data that contain mixed numerical and categorical features.

Beyond predictive performance, the integration of SHAP provides transparent and consistent explanations of model behavior, revealing that churn decisions are primarily driven by transactional and behavioral attributes rather than static demographic factors. This explainability component constitutes a key contribution of the proposed framework, as it enhances model auditability and supports transparency requirements in regulated banking environments. From an informatics perspective, this research contributes a methodologically coherent approach that integrates distribution-aware data generation, comparative ensemble modeling, and XAI, offering a reproducible framework for addressing imbalance and interpretability challenges in tabular banking data analytics.

Overall, the novelty introduced by this study is its unified framework that enhances data balancing through CTGAN-based augmentation, systematic boosting-based ensemble learning model comparison, and XAI through global feature attribution. This framework provides a coherent methodological foundation for robust and interpretable churn prediction in banking analytics. Future research may extend this approach to temporal and sequential transaction modeling, graph-based churn prediction for relational customer behavior, hybrid XAI, and real-time or adaptive learning mechanisms to support dynamic banking environments.

CONFLICT OF INTEREST

The authors affirm the absence of any competing interests related to the publication of this work. All analyses and interpretations were conducted objectively and without any commercial or personal affiliations with the capacity to exert undue influence on the research outcomes.

REFERENCES

[1] A. G. Văduva, S. V. Oprea, A. M. Niculae, A. Bâra, and A. I. Andreescu, "Improving Churn Detection in the Banking Sector: A Machine Learning Approach with Probability Calibration

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

Techniques," *Electronics (Switzerland)*, vol. 13, no. 22, Nov. 2024, doi: 10.3390/electronics13224527.

[2] A. Agnihotri and R. Saravanakumar, "Customer Retention in Banking: Utilizing AI and Machine Learning for Predictive Churn Analysis," in *Proceedings of 2025 3rd International Conference on Intelligent Systems, Advanced Computing, and Communication, ISACC 2025*, Institute of Electrical and Electronics Engineers Inc., 2025, pp. 140–144. doi: 10.1109/ISACC65211.2025.10969188.

[3] D. O. U. Orina, R. Rimiru, and W. Mwangi, "A Comparative Study of Predictive Data Mining Techniques for Customer Churn in the Banking Industry," in *1st International Conference of Intelligent Methods, Systems and Applications, IMSA 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 222–227. doi: 10.1109/IMSA58542.2023.10217514.

[4] A. Manzoor, M. Atif Qureshi, E. Kidney, and L. Longo, "A Review on Machine Learning Methods for Customer Churn Prediction and Recommendations for Business Practitioners," *IEEE Access*, vol. 12, pp. 70434–70463, 2024, doi: 10.1109/ACCESS.2024.3402092.

[5] U. Gani Joy, K. E. Hoque, M. Nazim Uddin, L. Chowdhury, and S. B. Park, "A Big Data-Driven Hybrid Model for Enhancing Streaming Service Customer Retention Through Churn Prediction Integrated With Explainable AI," *IEEE Access*, vol. 12, pp. 69130–69150, 2024, doi: 10.1109/ACCESS.2024.3401247.

[6] D. D. Ninditha Silalahi, Marsella, A. A. Valentino, I. S. Edbert, and D. Suhartono, "Bagging and Boosting for Predicting Bank Customer Churn," in *2023 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation, ICAMIMIA 2023 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ICAMIMIA60881.2023.10427686.

[7] S. C. K. Tékouabou, Ștefan C. Gherghina, H. Toulni, P. N. Mata, and J. M. Martins, "Towards Explainable Machine Learning for Bank Churn Prediction Using Data Balancing and Ensemble-Based Methods," *Mathematics*, vol. 10, no. 14, Jul. 2022, doi: 10.3390/math10142379.

[8] Y. Deng, D. Li, L. Yang, J. Tang, and J. Zhao, "Analysis and prediction of bank user churn based on ensemble learning algorithm," in *Proceedings of 2021 IEEE International Conference on Power Electronics, Computer Applications, ICPECA 2021*, Institute of Electrical and Electronics Engineers Inc., Jan. 2021, pp. 288–291. doi: 10.1109/ICPECA51329.2021.9362520.

[9] U. Mansoor, V. Sivakumar, and M. Jayabalan, "Customer Churn Prediction in The Banking Sector on Imbalance Dataset," in *International Conference on Integrated Intelligence and Communication Systems, ICIICS 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ICIICS59993.2023.10421738.

[10] J. Li, X. Bai, Q. Xu, and D. Yang, "Identification of Customer Churn Considering Difficult Case Mining," *Systems*, vol. 11, no. 7, Jul. 2023, doi: 10.3390/systems11070325.

[11] A. Soni, J. Mishra, and M. Dixit, "Comparative Study of Bank Customers Churn Prediction using AI/ML," in *IEEE International Conference on Communication Systems and Network Technologies*, Jabalpur: IEEE, 2024, pp. 1359–1365. doi: 10.1109/CSNT.2024.224.

[12] D. Hason Rudd, H. Huo, and G. Xu, "Improved Churn Causal Analysis Through Restrained High-Dimensional Feature Space Effects in Financial Institutions," *Human-Centric Intelligent Systems*, vol. 2, no. 3, pp. 70–80, Dec. 2022, doi: 10.1007/s44230-022-00006-y.

[13] A. S. Nair, A. Krishna, S. T. Gupta, and S. Susan, "Credit Card Fraud Detection using Soft Voting Ensemble with Imbalance Treatment," in *2025 5th International Conference on Intelligent Technologies, CONIT 2025*, Institute of Electrical and Electronics Engineers Inc., 2025. doi: 10.1109/CONIT65521.2025.11167470.

[14] I. N. M. Adiputra and P. Wanchai, "CTGAN-ENN: a tabular GAN-based hybrid sampling method for imbalanced and overlapped data in customer churn prediction," *J Big Data*, vol. 11, no. 1, Dec. 2024, doi: 10.1186/s40537-024-00982-x.

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

[15] A. Muneer, R. F. Ali, A. Alghamdi, S. M. Taib, A. Almaghthawi, and E. A. Abdullah Ghaleb, "Predicting customers churning in banking industry: A machine learning approach," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 26, no. 1, pp. 539–549, Apr. 2022, doi: 10.11591/ijeecs.v26.i1.pp539-549.

[16] B. A. Maulana and N. Hidayati, "Churn Prediction in Credit Customers Using Random Forest and XGBoost Methods," *Indonesian Journal of Data and Science*, vol. 6, no. 1, pp. 82–90, Mar. 2025, doi: 10.56705/ijodas.v6i1.215.

[17] A. Singh, R. Vashisth, N. Sindhwani, and G. Arora, "Credit Card Users Churn Prediction Using Ensemble Techniques," in *Proceedings - International Conference on Technological Advancements in Computational Sciences, ICTACS 2023*, Institute of Electrical and Electronics Engineers Inc., 2023, pp. 1283–1290. doi: 10.1109/ICTACS59847.2023.10390508.

[18] R. Bhuria *et al.*, "Ensemble-based customer churn prediction in banking: a voting classifier approach for improved client retention using demographic and behavioral data," *Discover Sustainability*, vol. 6, no. 1, Dec. 2025, doi: 10.1007/s43621-025-00807-8.

[19] A. O. Babatunde, S. A. Yinusa, I. D. Oladipo, and A. W. Asaju-Gbolagade, "View of Customer Churn Prediction in Neobanking System Using Predictive Analytics and Feature Selection," *Systems and Computing*, no. 1, pp. 27–43, 2025, doi: https://doi.org/10.64409/sycom.v1.i1.14.

[20] F. Sağlam and M. A. Cengiz, "A novel SMOTE-based resampling technique trough noise detection and the boosting procedure," *Expert Syst Appl*, vol. 200, Aug. 2022, doi: 10.1016/j.eswa.2022.117023.

[21] J. Černevičienė and A. Kabašinskas, "Explainable artificial intelligence (XAI) in finance: a systematic literature review," *Artif Intell Rev*, vol. 57, no. 8, Aug. 2024, doi: 10.1007/s10462-024-10854-8.

[22] O. Parise, R. Kronenberger, G. Parise, C. de Asmundis, S. Gelsomino, and M. La Meir, "CTGAN-driven synthetic data generation: A multidisciplinary, expert-guided approach (TIMA)," *Comput Methods Programs Biomed*, vol. 259, Feb. 2025, doi: 10.1016/j.cmpb.2024.108523.

[23] A. M. Salih *et al.*, "A Perspective on Explainable Artificial Intelligence Methods: SHAP and LIME," *Advanced Intelligent Systems*, vol. 7, no. 1, Jan. 2025, doi: 10.1002/aisy.202400304.

[24] H. A. Raouf, M. M. Fouda, and M. I. Ibrahem, "Revolutionizing User Authentication Exploiting Explainable AI and CTGAN-Based Keystroke Dynamics," *IEEE Open Journal of the Computer Society*, vol. 6, pp. 97–108, 2025, doi: 10.1109/OJCS.2024.3513895.

[25] A. Alzahrani, "Early Detection of Lung Cancer Using Predictive Modeling Incorporating CTGAN Features and Tree-Based Learning," *IEEE Access*, vol. 13, pp. 34321–34333, 2025, doi: 10.1109/ACCESS.2025.3543215.

[26] O. Habibi, M. Chemmakha, and M. Lazaar, "Imbalanced tabular data modelization using CTGAN and machine learning to improve IoT Botnet attacks detection," *Eng Appl Artif Intell*, vol. 118, Feb. 2023, doi: 10.1016/j.engappai.2022.105669.

[27] C. G. L. Pringandana and K. Kusnawi, "A Comparative Analysis of Hyperparameter-Tuned XGBoost and LightGBM for Multiclass Rainfall Classification in Jakarta," *Jurnal Teknik Informatika (Jutif)*, vol. 6, no. 4, pp. 2467–2483, Aug. 2025, doi: 10.52436/1.jutif.2025.6.4.4965.

[28] I. Maulana, A. M. Siregar, S. A. P. Lestari, and S. Faisal, "OPTIMAL STUDY OF REAL-ESTATE PRICE PREDICTION MODELS USING MACHINE LEARNING," *Jurnal Teknik Informatika (Jutif)*, vol. 5, no. 4, pp. 1149–1164, Aug. 2024, doi: 10.52436/1.jutif.2024.5.4.2565.

[29] E. E. Başakın, Ö. Ekmekcioğlu, P. C. Stoy, and M. Özger, "Estimation of daily reference evapotranspiration by hybrid singular spectrum analysis-based stochastic gradient boosting," *MethodsX*, vol. 10, 2023, doi: 10.1016/j.mex.2023.102163.

[30] C. Yu, Y. Jin, Q. Xing, Y. Zhang, S. Guo, and S. Meng, "Advanced User Credit Risk Prediction Model Using LightGBM, XGBoost and Tabnet with SMOTEENN," in *2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems, ICPICS 2024*, Institute

Vol. 7, No. 3, June 2026, Page. 2376-2394
https://jutif.if.unsoed.ac.id
DOI: https://doi.org/10.52436/1.jutif.2026.7.3.5578

of Electrical and Electronics Engineers Inc., 2024, pp. 876–883. doi: 10.1109/ICPICS62053.2024.10796247.

[31] R. K. Makumbura *et al.*, "Advancing water quality assessment and prediction using machine learning models, coupled with explainable artificial intelligence (XAI) techniques like shapley additive explanations (SHAP) for interpreting the black-box nature," *Results in Engineering*, vol. 23, Sep. 2024, doi: 10.1016/j.rineng.2024.102831.