ENHANCING LOAN APPROVAL DECISION-MAKING: AN INTERPRETABLE MACHINE LEARNING APPROACH USING LIGHTGBM FOR DIGITAL ECONOMY DEVELOPMENT

Teuku Rizky Noviandy¹*, Ghalieb Mutig Idroes² and Irsan Hardi³

¹Department of Informatics, Faculty of Mathematics and Natural Sciences, Universitas Syiah Kuala, Banda Aceh 23111, Indonesia
²Energy and Green Economics Unit, Graha Primera Saintifika, Aceh Besar 23371, Indonesia
³Economic Modeling and Data Analytics Unit, Graha Primera Saintifika, Aceh Besar 23371, Indonesia

¹trizkynoviandy@gmail.com, ²ghaliebidroes@outlook.com ³irsan.hardi@gmail.com

ABSTRACT

This study aims to enhance loan approval decision-making in the digital economy using an interpretable machine learning approach. The primary research question investigates how integrating an interpretable machine learning approach can improve the accuracy and transparency of loan approval processes. We employed LightGBM, a gradient-boosting framework for loan approval classification, optimized via Random Search hyperparameter tuning and validated using 10-fold cross-validation. We incorporated the Shapley Additive exPlanations (SHAP) framework to address the challenge of interpretability in machine learning. The LightGBM model outperformed conventional algorithms (Decision Tree, Random Forest, AdaBoost, and Extra Trees) in accuracy (98.13%), precision (97.78%), recall (97.17%), and F1-score (97.48%). The study demonstrates that using an interpretable machine learning approach with LightGBM and SHAP can significantly improve the accuracy and transparency of loan approval decisions. This method offers a promising avenue for financial institutions to enhance their loan approval mechanisms, ensuring more reliable, efficient, and transparent decision-making in the digital economy. The study also underscores the importance of interpretability in deploying machine learning solutions in sectors with significant socio-economic impacts.

Keywords: Artificial Intelligence, Light Gradient Boosting Machine, Machine Learning, SHAP

1. Introduction

Loans are pivotal in facilitating financial transactions and enabling individuals and businesses to realize their aspirations (Kariv & Coleman, 2015). Whether for purchasing a home, expanding a business, or covering unexpected expenses, loans are integral to economic growth and personal advancement (Makinde, 2016; Saiti & Trenovski, 2022). However, the process of
approving loans involves intricate decision-making, where financial institutions evaluate numerous factors to determine the creditworthiness of applicants. As the demand for loans continues to rise, the need for effective and efficient loan approval mechanisms becomes increasingly crucial (Dansana et al., 2023).

Traditionally, the loan approval process has been characterized by manual assessment methods that rely heavily on historical data and rigid criteria (Tchakoute Tchuigoua, 2018). These traditional approaches often struggle to adapt to the evolving landscape of financial dynamics, resulting in inefficiencies and suboptimal decision-making. The limitations of these methods become particularly evident when faced with complex and dynamic economic conditions, leading to delays, inaccuracies, and, sometimes, overlooking potentially creditworthy applicants (Lamichhane, 2022; Purificato et al., 2023).

In recent years, machine learning has brought about a paradigm shift in the domain of loan approval. Various studies have explored the application of machine learning algorithms to enhance the accuracy and efficiency of decision-making processes (Alaradi & Hilal, 2020; Orji et al., 2022; Sheikh et al., 2020). A widely adopted machine learning framework known for its efficiency and effectiveness is the Light Gradient Boosting Machine (LightGBM) (Ke et al., 2017). Its notable strengths include the capability to handle large datasets and intricate feature interactions, resulting in high predictive accuracy (Ponsam et al., 2021). LightGBM has proven its performance in various applications, for instance, in energy management, drug discovery, and roadway safety (Musbah et al., 2022; Noviandy et al., 2023; Wen et al., 2021).

Despite the promising strides made in this direction, a notable gap exists in understanding and interpreting the decision rationale of these complex models. The lack of transparency poses a significant challenge as stakeholders, including applicants and regulatory bodies, seek a clearer understanding of the factors influencing loan approval outcomes (Purificato et al., 2023). This lack of transparency undermines trust in the lending process and hampers efforts to identify and mitigate potential biases or errors in the decision-making process (Brotcke, 2022). Thus, there is a pressing need for greater transparency and accountability in developing and deploying these models to ensure fairness and reliability in lending practices.

Recognizing the importance of transparency and interpretability in the loan approval process, this study emphasizes the need for an interpretable approach. Interpretable models not only enhance the trustworthiness of the decision-making process but also empower stakeholders with insights into the inner workings of the model (Cakiroglu et al., 2024). In this context, the study leverages the Shapley Additive exPlanations (SHAP) framework, a powerful tool for explaining the output of machine learning models (Lundberg & Lee, 2017). By incorporating SHAP interpretability into the loan approval process, the aim is to provide a clearer understanding of the factors influencing decisions, bridging the gap between the complexity of machine learning models and the need for transparency.

2. Methodology
2.1 Data Acquisition and Preprocessing

The dataset used in this study is sourced from Kaggle, a renowned platform for machine learning datasets (Sharma, 2023). The dataset comprises 11 essential features related to loan applications, with the target variable being the binary classification of loan approval status (approved or rejected). The dataset consists of 4,269 instances, with 2,656 cases marked as approved and 1,613 cases marked as rejected. The data type and description for each feature are described in Table 1.

A thorough preprocessing stage ensures the dataset is well-prepared for model training. This involves a meticulous examination of missing values in any features, and if identified, appropriate strategies such as imputation or removal are applied to handle the missing data (Hui...
et al., 2023; Idroes et al., 2021). Categorical features undergo one-hot encoding to convert them into a format suitable for machine learning algorithms. Additionally, the dataset is divided into training and testing sets, with 80% allocated for training and 20% for testing. This strategic partitioning ensures the model is trained on a substantial portion of the data while maintaining a separate and independent set for an unbiased evaluation of its performance (Noviandy et al., 2023b).

Table 1. Description of the Features

<table>
<thead>
<tr>
<th>No.</th>
<th>Feature Name</th>
<th>Data Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>no_of_dependents</td>
<td>Numeric</td>
<td>Number of dependents</td>
</tr>
<tr>
<td>2</td>
<td>education</td>
<td>Categorical</td>
<td>The education level of the applicant</td>
</tr>
<tr>
<td>3</td>
<td>self_employed</td>
<td>Categorical</td>
<td>Employment status (self-employed or not)</td>
</tr>
<tr>
<td>4</td>
<td>income_annum</td>
<td>Numeric</td>
<td>Annual income of the applicant</td>
</tr>
<tr>
<td>5</td>
<td>loan_amount</td>
<td>Numeric</td>
<td>Requested loan amount</td>
</tr>
<tr>
<td>6</td>
<td>loan_term</td>
<td>Numeric</td>
<td>Duration of the loan in months</td>
</tr>
<tr>
<td>7</td>
<td>cibil_score</td>
<td>Numeric</td>
<td>Credit score of the applicant</td>
</tr>
<tr>
<td>8</td>
<td>residential_assets_value</td>
<td>Numeric</td>
<td>Value of residential assets</td>
</tr>
<tr>
<td>9</td>
<td>commercial_assets_value</td>
<td>Numeric</td>
<td>Value of commercial assets</td>
</tr>
<tr>
<td>10</td>
<td>luxury_assets_value</td>
<td>Numeric</td>
<td>Value of luxury assets</td>
</tr>
<tr>
<td>11</td>
<td>bank_asset_value</td>
<td>Numeric</td>
<td>Value of assets held in bank</td>
</tr>
</tbody>
</table>

2.2 LightGBM Model

This study employed LightGBM, a gradient-boosting framework, for loan approval classification. LightGBM was chosen for its efficiency, effectiveness in handling large datasets, and ability to capture non-linear relationships between features and target variables (Noviandy et al., 2023d; Sevgen & Abdikan, 2023). Additionally, LightGBM's speed and scalability make it well-suited for the large-scale datasets commonly encountered in finance and banking.

To optimize the LightGBM model, a random search hyperparameter tuning approach was conducted (Mantovani et al., 2015). The random search technique explores various combinations from a predefined search space to identify the set of hyperparameters that maximizes the model's predictive performance (Idroes et al., 2023). The hyperparameter tuning process aims to find the optimal combination of these parameters, enhancing the LightGBM model's ability to accurately predict loan approval outcomes. The hyperparameters subjected to tuning are detailed in Table 2.

Table 2. Hyperparameter Space Used for Tuning

<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Distribution/Range</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>n_estimators</td>
<td>Discrete Uniform (50, 500)</td>
<td>Number of boosting rounds</td>
</tr>
<tr>
<td>max_depth</td>
<td>Discrete Uniform (3, 15)</td>
<td>Maximum depth of trees</td>
</tr>
<tr>
<td>learning_rate</td>
<td>Linear Space (0.01, 0.3), 100 steps</td>
<td>Learning rate</td>
</tr>
<tr>
<td>subsample</td>
<td>Linear Space (0.2, 1.0), 100 steps</td>
<td>Subsample ratio</td>
</tr>
<tr>
<td>colsample_bytree</td>
<td>Linear Space (0.2, 1.0), 100 steps</td>
<td>Feature subsample ratio</td>
</tr>
</tbody>
</table>

To ensure the robustness and generalizability of the model, a 10-fold cross-validation strategy was implemented during the training process (Maulana et al., 2023). Cross-validation
is a technique used to assess the model's ability to generalize to unseen data. In 10-fold cross-validation, the dataset is divided into ten equal-sized subsets or folds. The model is trained on nine folds and validated on the remaining fold (Noviandy et al., 2023a). This process is repeated ten times, with each fold representing the validation set once. The performance metrics obtained from each fold are then averaged to estimate the model's overall performance. Cross-validation reduces the risk of overfitting, and the model's reliability and generalizability are enhanced (Berrar, 2019).

2.3 Model Evaluation and Interpretation

The performance of the approach was compared against four well-established machine learning algorithms: Decision Tree, Random Forest, Adaptive Boosting (AdaBoost), and Extra Trees. This benchmark evaluation provided insights into how the model fares in comparison to widely used classical techniques (Mansur Huang et al., 2021; Shafie et al., 2023).

The model's performance is evaluated using four essential metrics: accuracy, which measures the overall correctness of the model's predictions; precision, indicating the ratio of correctly predicted positive observations to the total predicted positives; recall, representing the ratio of correctly predicted positive observations to the actual positives; and F1-score, a harmonic mean of precision and recall, offering a balanced assessment of the model's effectiveness (Noviandy et al., 2023e). The predictions on the test set undergo a meticulous examination through a confusion matrix, providing insights into the specific types of errors made by the model, such as false positives and false negatives (Lee & Jemain, 2019). The accuracy, precision, recall, and F1-score equations are presented in Equations (1) - (4), respectively.

\[
\text{Accuracy} = \frac{TP + FN}{FP + FN + TP + TN} \tag{1}
\]

\[
\text{Precision} = \frac{TP}{TP + FP} \tag{2}
\]

\[
\text{Recall} = \frac{TP}{FN + TP} \tag{3}
\]

\[
F1 - \text{Score} = 2 \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}
\]

TP (True Positive) represents instances where the model correctly predicts and identifies loan rejections, FN (False Negative) denotes instances where the model incorrectly predicts loan approval when, in reality, the loan is rejected, FP (False Positive) indicates instances where the model wrongly predicts loan rejection when the loan is approved, and TN (True Negative) represents cases where the model accurately predicts and identifies approved loans (Suhendra et al., 2023).

To further assess the model's ability to discriminate between classes, a Receiver Operating Characteristic (ROC) curve is generated. This curve illustrates the trade-off between the true and false positive rates at various classification thresholds (Supriatna et al., 2023). The Area Under the Curve (AUC) score quantifies the ROC curve's performance, with a higher AUC indicating superior discriminative power. These comprehensive evaluation methods offer a nuanced understanding of the model's classification capabilities and highlight areas for potential improvement (Noviandy et al., 2023c).
Additionally, model interpretation is enhanced through the use of SHAP, which provides valuable insights into the importance of features and the contribution of each feature to individual predictions. SHAP values offer a nuanced understanding of the factors influencing the model's decisions, shedding light on each feature's specific role in the overall prediction (Marcílio & Eler, 2020). This interpretability tool allows stakeholders to grasp the rationale behind the model's outputs, fostering transparency and trust in its decision-making process (Le et al., 2022).

3. Results and Discussion

3.1 Models Performance

The LightGBM model employed in this study underwent optimization using random search hyperparameter tuning, resulting in the identification of the best-performing hyperparameters: 'colsample_bytree': 0.97, 'learning_rate': 0.265, 'max_depth': 11, 'n_estimators': 229, 'subsample': 0.70. These hyperparameters collectively define the model's architecture, influencing its ability to capture complex patterns and relationships within the loan approval dataset. Notably, a high 'colsample_bytree' value indicates that the model samples many features when constructing each tree, enhancing its diversity and robustness. The 'learning_rate' and 'n_estimators' parameters control the trade-off between model accuracy and training speed, with the specified values suggesting a careful balance for optimal performance. The 'max_depth' parameter signifies the maximum depth of each tree, influencing the model's capacity to capture intricate decision boundaries. Lastly, the 'subsample' parameter determines the fraction of data randomly selected for each boosting round, contributing to the model's adaptability to different subsets of the dataset.

Compared to other machine learning models, the LightGBM model outperformed the Decision Tree, Random Forest, AdaBoost, and Extra Trees models across all metrics. The comprehensive results in Table 3 illustrate the LightGBM model's robustness and suitability for the task of loan approval, showcasing superior performance in terms of reliability, precision, and effectiveness in a critical financial domain. The LightGBM model achieved an impressive accuracy of 98.13%, indicating high overall correctness in its predictions. This metric is particularly significant in loan approval, reflecting the model's ability to classify approved and rejected loan applications correctly. Precision, at 97.78%, suggests that the model is highly reliable in its positive predictions, meaning that when it predicts loan rejection, it is correct most of the time. This is crucial for minimizing the risk of approving unqualified applicants. The recall of 97.17% indicates the model's proficiency in identifying true positive cases, the correctly identified rejected loans. A high recall is essential to ensure that deserving applicants are not mistakenly approved. Finally, the F1-Score, a balanced measure of precision and recall, stands at 97.48%. This score further underscores the model's effectiveness in providing reliable and balanced decision-making capabilities. These metrics collectively demonstrate the LightGBM model's robustness and suitability for loan approval, offering a high degree of reliability and precision in a critical financial domain.

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy (%)</th>
<th>Precision (%)</th>
<th>Recall (%)</th>
<th>F1-score (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LightGBM</td>
<td>98.13</td>
<td>97.78</td>
<td>97.17</td>
<td>97.48</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>97.07</td>
<td>96.81</td>
<td>95.28</td>
<td>96.04</td>
</tr>
<tr>
<td>Random Forest</td>
<td>97.89</td>
<td>97.77</td>
<td>96.54</td>
<td>97.15</td>
</tr>
<tr>
<td>AdaBoost</td>
<td>96.72</td>
<td>96.18</td>
<td>94.97</td>
<td>95.57</td>
</tr>
<tr>
<td>Extra Trees</td>
<td>95.20</td>
<td>93.42</td>
<td>93.71</td>
<td>93.56</td>
</tr>
</tbody>
</table>
The confusion matrix presented in Table 4 clearly depicts the LightGBM model's predictive accuracy on the test set. The matrix shows that the model correctly predicted 309 rejections and 529 approvals, which are true negatives and true positives, respectively. These numbers indicate the model's ability to classify both classes—rejected and approved loans correctly. However, there are 16 instances where the model's predictions deviated from the actual outcomes. Specifically, the model incorrectly predicted 9 cases as approved when they were rejected and 7 cases as rejected when they were approved. Compared to predictions by other models, LightGBM consistently exhibited higher accuracy, showcasing its robust performance in minimizing both false positive and false negative predictions, thus enhancing its overall reliability in the loan approval process.

<table>
<thead>
<tr>
<th>Model</th>
<th>Actual</th>
<th>Rejected</th>
<th>Approved</th>
</tr>
</thead>
<tbody>
<tr>
<td>LightGBM</td>
<td>Rejected</td>
<td>309</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Approved</td>
<td>7</td>
<td>529</td>
</tr>
<tr>
<td>Decision Tree</td>
<td>Rejected</td>
<td>303</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>Approved</td>
<td>15</td>
<td>526</td>
</tr>
<tr>
<td>Random Forest</td>
<td>Rejected</td>
<td>307</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Approved</td>
<td>7</td>
<td>529</td>
</tr>
<tr>
<td>AdaBoost</td>
<td>Rejected</td>
<td>302</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Approved</td>
<td>12</td>
<td>524</td>
</tr>
<tr>
<td>Extra Trees</td>
<td>Rejected</td>
<td>298</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Approved</td>
<td>20</td>
<td>515</td>
</tr>
</tbody>
</table>

### 3.2 Interpretation Using SHAP

SHAP values were employed to get deeper insights into the decision-making process of the LightGBM model, as depicted in Figure 1. The SHAP values provide a measure of the impact of each feature on the model's prediction, offering an interpretation of the model's behavior. From the SHAP bar plot, it is evident that the cibil_score is the most influential feature, with a mean SHAP value significantly higher than the rest, emphasizing its critical role in determining loan approval outcomes. This indicates that an applicant's credit score is the most substantial factor in the model's assessment, heavily influencing the prediction of approval or rejection. Other features such as loan_term, loan_amount, and income_annum also contribute to the decision, albeit to a lesser extent, suggesting that the model considers a blend of an applicant's financial stability and the requested loan's characteristics when making a prediction. Interestingly, features like residential_assets_value have some impact. However, factors such as education, self-employed status, and the number of dependents appear to have little to no direct influence on the model's predictions. This interpretation, enabled by SHAP values, enhances transparency in the model's predictive behavior and provides stakeholders with valuable insights into which factors are most pertinent for loan approval decisions.
The SHAP beeswarm plot presented in Figure 2 offers a more granular view of the contribution of each feature to the LightGBM model’s predictions compared to the bar plot previously discussed. Unlike the bar plot, which shows the mean impact of the features, the beeswarm plot provides a distribution of the SHAP values for each feature for individual predictions. Each point on the beeswarm plot represents a SHAP value for an individual instance in the dataset, indicating how much each feature pushed the model's output from the base value (the average model output over the dataset) to the actual model output. For example, the cibil_score feature shows a wide spread of SHAP values, with a dense collection of points on the higher end. In many instances, a high cibil_score significantly increases the likelihood of loan approval. Conversely, lower scores contribute to a lower predicted probability of approval. The distribution of points for loan_term, loan_amount, and income_annum also reveals variability in their impact on the model's output, with positive and negative contributions, reflecting the nuanced relationship between these features and the loan approval decision. The beeswarm plot thus provides a detailed look at the heterogeneity in the impact of features across different predictions, highlighting the model's complex decision-making process. By revealing this variance, stakeholders can understand which features are important on average and how the model assesses individual instances. This depth of insight is crucial for interpreting the model's behavior in specific cases, fostering greater trust and transparency in its use for financial decision-making.
3.3 Implications, Limitations, and Future Research

The findings of this study have significant implications that span across various stakeholders within the financial services industry, including financial institutions, loan applicants, regulatory bodies, and the data science community. By offering valuable insights into the potential advantages and best practices for integrating advanced machine learning models, such as LightGBM, this study provides a comprehensive guide for decision-makers, enabling them to make well-informed choices regarding adopting and implementing these cutting-edge technologies. The study's findings serve as a roadmap for navigating the complex landscape of artificial intelligence (AI) integration in the financial sector, ensuring that the benefits are maximized while potential risks are mitigated.

For financial institutions, incorporating sophisticated machine learning models like LightGBM offers an opportunity to revolutionize loan approval processes, increasing their accuracy and efficiency. This transformation could lead to substantial cost reductions, lower default rates, and improved customer experiences through faster service delivery. Moreover, loan applicants stand to benefit from greater transparency and fairness in loan determinations, potentially gaining insights into the factors influencing their creditworthiness as illuminated by SHAP interpretability.

Regulatory bodies can also draw valuable insights from this study. The methodology presented offers a blueprint for the effective and transparent adoption of AI, aiding in the enforcement of fair lending laws and fostering trust in automated financial systems. By encouraging interpretable models, regulators can ensure that financial institutions remain accountable and that lending practices are free from bias.

Furthermore, this study serves as a compelling case study for the data science community, highlighting the critical importance of model interpretability, particularly in sectors with substantial socio-economic impacts. As data scientists and researchers continue to develop and refine machine learning models, prioritizing interpretability alongside performance will be essential to ensure AI's responsible and ethical deployment in high-stakes domains like financial services.

However, it is also important to consider the ethical implications of relying on machine learning for loan approvals. While the model performs well statistically, the financial industry must ensure that it does not perpetuate existing biases or unfair practices. The apparent minimal influence of features such as education and employment status on the model's decisions raises questions about the comprehensiveness of the data and the potential for unintended bias. It is important that models are regularly audited for fairness and incorporate a wide range of socioeconomic factors to make equitable decisions. Additionally, the current study relies on a dataset that may not capture the full spectrum of variables influencing creditworthiness. The exclusion or underrepresentation of certain variables could skew the model's predictions. Furthermore, while the model has been validated statistically, its real-world applicability across different demographic groups and economic contexts has yet to be established. Future studies could address these limitations by including more diverse variables, testing the model across various populations, and developing more user-friendly interpretability interfaces.

4. Conclusions

The approach presented in this study offers a robust and interpretable framework that could be instrumental in shaping the future of financial decision-making. It fosters a more inclusive, efficient, and transparent financial landscape, aligning with the evolving demands of the digital economy. However, continual evaluation and adaptation of such models are essential to ensure their relevance and fairness, considering the dynamic nature of economic and social factors influencing financial decisions.
Acknowledgement

We would like to express our sincere gratitude to our respective institutions for their unwavering support throughout the duration of this project.

Funding

This study received no external funding or financial support.

Author Contribution

Teuku Rizky Noviandy: Conceptualization, Methodology, Software, Writing-Original draft preparation, Writing-Reviewing and Editing. Ghalieb Mutig Idroes: Methodology, Visualization, Investigation, Writing-Reviewing and Editing. Irsan Hardi: Conceptualization, Validation, Writing-Original draft preparation.

Conflict of Interest

The authors declare no conflict of interest.

References


