Available online at website: https://jurnal.iaii.or.id/index.php/RESTI

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 9 No. 5 (2025) 983 - 990 e-ISSN: 2580-0760

Explainable Ensemble Learning Framework with SMOTE, SHAP and LIME for Predicting 30-Day Readmission in Diabetic Patients

Joshua Pinem[1], Widi Astuti[2], Adiwijaya[3]
[1,2,3]Informatics Study Program, School of Computing, Telkom University, Bandung, Indonesia
[1]jopisan@student.telkomuniversity.ac.id, [2]widiwdu@telkomuniversity.ac.id, [3]adiwijaya@telkomuniversity.ac.id

Abstract

Hospital readmission among diabetic patients poses a significant burden on healthcare systems due to its frequency and associated costs. This study presents a machine learning framework for predicting 30-day readmission in diabetic patients using the Diabetes 130-US Hospitals dataset. The framework integrates data preprocessing, SMOTE for class balancing, ensemble learning, and explainable AI (SHAP and LIME) to enhance both accuracy and interpretability. Multiple models were evaluated, and the best performance was achieved by a weighted ensemble with a recall of 89.43% and an F1-score of 0.6612, indicating strong sensitivity. Explainability analysis using SHAP and LIME highlighted key predictors, notably Medication Change Status and Inpatient Admissions, which are clinically relevant. By combining predictive performance with transparent explanations, the proposed framework offers a practical and trustworthy tool for clinical decision support in managing diabetic readmissions.

Keywords: diabetes; ensemble learning; explainable AI; machine learning; readmission prediction; SMOTE

How to Cite: J. Pinem, W. Astuti, and Adiwijaya, "Explainable Ensemble Learning Framework with SMOTE, SHAP and LIME for Predicting 30-Day Readmission in Diabetic Patients ", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 9, no. 5, pp. 983 - 990, Sep. 2025.
Permalink/DOI: https://doi.org/10.29207/resti.v9i5.6977

Received: July 19, 2025
Accepted: August 25, 2025
Available Online: September 29, 2025

This is an open-access article under the CC BY 4.0 License
Published by Ikatan Ahli Informatika Indonesia

1. Introduction

Diabetes is a major global health challenge, affecting over 415 million individuals worldwide [1], and was listed among the top ten causes of death in 2021 [2]. Its management requires continuous care to maintain stable blood glucose and prevent complications [3]. contributing to the burden on healthcare systems, particularly from high readmission rates in diabetic patients [4], [5]. The 30-day readmission rate has become a key performance indicator, as used by the Centers for Medicare and Medicaid Services [6], and reflects both the quality of patient care and its financial impact on the healthcare system [7].

The 30-day readmission rate for diabetic patients ranges from 14.4% to 22.7%, which is much higher than the readmission rate for the entire patient population, which ranges from 8.5% to 13.5% [8]. From an economic perspective, the cost burden associated with readmission is substantial. In 2012, total hospital care expenditures for diabetes patients in the United States reached $124 billion, with approximately $25 billion (20%) of that amount specifically allocated to managing 30-day readmissions [9].

Reducing readmissions within 30 days in diabetic patients can significantly lower healthcare costs and enhance the quality of care. A simple 5% reduction in the readmission rate would result in 82,754 fewer admissions based on the Agency for Healthcare Research and Quality Nationwide Inpatient Survey from 2012, which translates to annual cost savings of $1.2 billion [10]. Given this urgency, it is essential to accurately predict the risk of hospital readmission in diabetic patients to improve cost efficiency and the overall quality of healthcare services. In this context, Machine Learning (ML) offers an effective tool for detecting readmission risks in diabetic patients due to its ability to identify complex patterns within large-scale clinical data [11].

A key challenge in developing Machine Learning (ML) models for healthcare is their black-box nature, where the decision-making process is not easily interpretable [12]. This lack of transparency hinders understanding of which features drive predictions and poses a barrier to

clinical adoption, as clinicians need clear, justifiable reasoning for each decision. Explainable Artificial Intelligence (XAI) addresses this by revealing the contribution of individual features to model predictions [13]. Besides interpretability, healthcare ML models face complex datasets with noise, class imbalance, and missing values, which reduce accuracy and generalization. Effective preprocessing is therefore crucial [14], as shown by Ramadhan et al. [15], who showed that careful preprocessing significantly improved Random Forest performance in predicting type 2 diabetes.To address the challenge of predicting 30-day readmission in diabetic patients, several studies have proposed machine learning models that emphasize both accuracy and model transparency. For instance, Mpofu et al. [16] applied the CRISP-DM framework using Random Forest and Logistic Regression, achieving 88.97% accuracy. However, their approach lacked interpretability, did not address class imbalance, and was limited in algorithm diversity, factors that are crucial for clinical applicability and model robustness.

Emi-Johnson and Nkrumah [4] used XGBoost with SHAP to predict 30-day readmission in diabetic patients, achieving an AUC-ROC of 0.667. However, they overlooked class imbalance—an important limitation that may reduce sensitivity in detecting high-risk, minority-class patients.

Lu and Uddin [17] used ensemble models with preprocessing and random under-sampling for class imbalance to predict hospital readmissions. They applied LIME for interpretability but achieved modest accuracy (68.63%) and lost majority class information due to under-sampling. The study did not combine SHAP and LIME, limiting explanation quality. Similarly, Shang et al. [7] used SMOTE and feature engineering with random forest for 30-day diabetic readmission prediction. Despite good performance, their model lacked explainability, affecting clinical trust.

Alnazari et al. [18] predicted 30-day readmission after kidney transplantation using Gradient Boosting. Their framework combined preprocessing, model validation, and dual explainability via SHAP and LIME. Although the study achieved an accuracy of 79.6%, it was limited to post-transplant patients and not focused on diabetes-related readmissions. Still, their emphasis on transparency via dual-XAI serves as a precedent that supports the direction of this study.

Zhang et al. [19] also conducted a study predicting readmission 30 days after hospitalization for patients with Acute Heart Failure (AHF). This study performed preprocessing using KNN for missing value imputation and recursive feature elimination for feature selection. A comparison was made between five machine learning algorithms. This study also applied XAI using SHAP. The best results were obtained from the XGBoost model, with an AUC of 0.763. However, this study did not include the clinical dynamics of patients during hospitalization.

Gao et al. [20] proposed an interpretable machine learning framework for predicting 30-day hospital readmission using a continuous regression tree. Their method followed a two-step process, starting with the training of a black-box neural network model and then extracting a regression tree from its output scores, using synthetically generated data to approximate the model's decision boundaries. While this approach offered a novel way to enhance interpretability, the AUC achieved was relatively modest at 0.711. In addition, the study did not utilize widely accepted explainability techniques, such as SHAP or LIME, which limits its alignment with current best practices in interpretable machine learning, particularly in clinical settings.

Martin et al. [21] predicted unscheduled hospital readmission (UHR) rates using logistic regression, followed by comparisons with other machine learning algorithms. The study included preprocessing steps such as numerical standardization and categorical encoding to handle clinical data effectively. For interpretability, both SHAP and LIME were employed, enabling insights into global and local feature importance. However, the study did not specifically focus on diabetic patients or address class imbalance, limiting its applicability in specialized clinical contexts.

Tasin et al. [22] developed a diabetes prediction system by combining local and global data and applying data balancing techniques such as SMOTE and ADASYN. The model also uses Explainable AI methods, including SHAP for global interpretation and LIME for local interpretation, which enhance transparency and clinical trust. Although the study focused on diabetes prediction rather than readmission, this integrated approach of preprocessing, balancing, and dual explainability offers a valuable reference for developing accurate and explainable health prediction models.

Despite advances in hospital readmission prediction, most prior studies have not presented clinically relevant pipelines integrating robust preprocessing, effective class balancing, diverse ensemble strategies, and comprehensive explainability. Many focus solely on performance without transparency, while others apply interpretability but overlook class imbalance, limiting model generalization. Automated, model-agnostic methods such as SHAP and LIME are more suitable than model-specific techniques because they apply across algorithms and provide complementary perspectives. SHAP offers consistent global feature importance, while LIME explains individual predictions intuitively, both valuable in clinical contexts. Addressing this gaps, this study make three key contributions: (1) developing a pipeline integrating systematic preprocessing, SMOTE-based class balancing, and multiple ensemble models to enhance sensitivity; (2) applying recall-focused evaluation to prioritize detection of high-risk patients; and (3) incorporating dual SHAP and LIME explainability to ensure transparency and strengthen clinical trust.

2. Methods

2.1 Research Approach

This section outlines the research methodology adopted in this study. The methodological design provides a structured process that begins with the preparation of data and continues through a series of steps to develop and validate the predictive framework. Each stage plays an important role in ensuring that the resulting model is both reliable and suitable for application in a clinical context. An overview of the complete framework is illustrated in Figure 1.

To address the challenge of predicting 30-day readmission in diabetic patients, this study applies a machine learning based classification approach in a structured manner, including preprocessing, class distribution balancing, and the use of ensemble learning algorithms. Ensemble methods are selected because they improve predictive accuracy and reduce overfitting by combining diverse base models. In this study, six base learner models are integrated through soft voting to enhance stability and generalization. To ensure the results are interpretable and clinically meaningful, Explainable AI methods such as SHAP and LIME are applied to the best performing model, providing both global and local explanations that strengthen transparency and clinical trust.

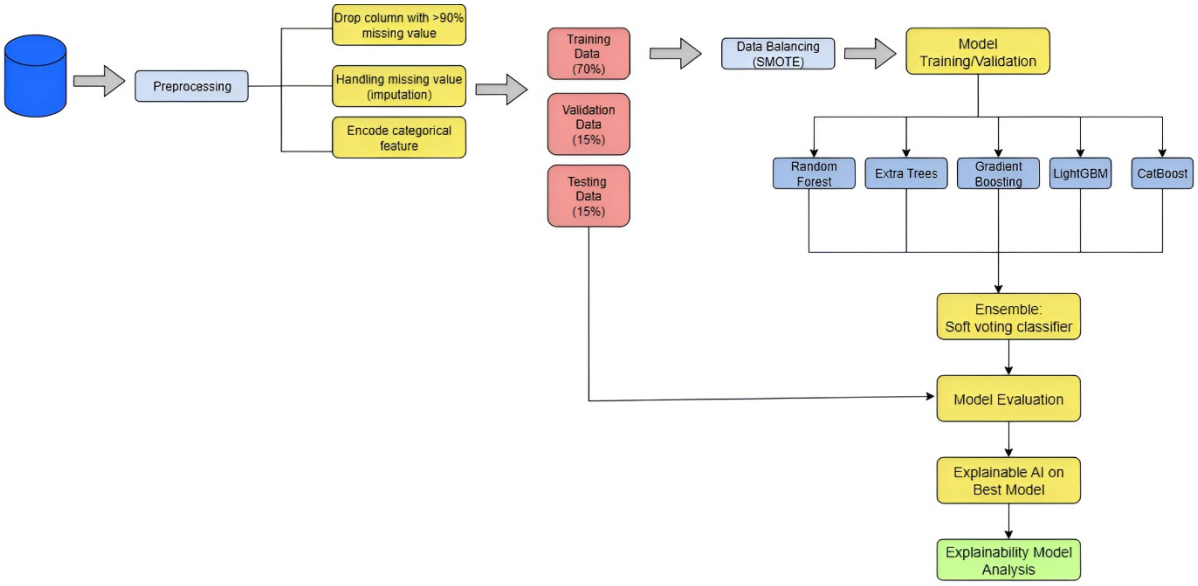


Figure 1. Overview of the readmission prediction framework

2.2 Dataset Description

This study utilizes the Diabetes 130-US Hospitals dataset, which contains clinical data collected over a ten-year period (1999–2008) from 130 hospitals across the United States [23]. It comprises over 100,000 inpatient records of individuals diagnosed with diabetes. Each entry corresponds to a single hospital admission and includes detailed information such as demographic data (age, gender, race), laboratory results, treatment types, ICD-9 diagnostic codes, and length of stay (up to 14 days). Additionally, the dataset records discharge outcomes and indicate whether the patient was readmitted. A small snippet of the dataset is presented in Table 1, showing several key attributes such as encounter ID, age, gender, weight, and readmission status. The symbol "?" in the Weight column denotes missing values. The original dataset contains over 50 features, but only a subset is presented here for clarity. This dataset is publicly accessible through the UCI Machine Learning Repository(https://archive.ics.uci.edu/ml/datasets/Diabetes+130-US+hospitals+for+years+1999-2008).

This dataset is real clinical data that has quality challenges, such as high missing values and class distribution imbalance. One extreme example is the Weight attribute, which has 97% missing values. In addition, the target variable Readmitted shows a very unbalanced distribution, with only about 11% of patients experiencing readmission within 30 days, while the remaining 88% did not experience readmission. This imbalance can cause bias in model training toward the majority class [24].

Table 1. Snippet of the dataset

encounterid	Age	Gender	Weight	Readmitted
2278392	[0-10)	Female	?	NO
149190	[10-20)	Female	?	>30
64410	[20-30)	Female	?	NO
500364	[30-40)	Male	?	NO

2.3. Data Preprocessing

During the preprocessing stage, the percentage of missing values in each feature is checked. Specifically, features with more than 90% missing values are removed from the dataset. Features with very high missing values have very limited information and risk causing bias in model training [25].

For features with less than 90% missing values, imputation will be performed. Missing values in features use mean imputation while missing values in categorical features use mode imputation [3]. Categorical features are then encoded, and standardized data can improve model performance [11].

2.3. Balancing Class

After preprocessing, class balancing was performed on the readmitted variable, which showed a substantial imbalance with only about 11% of patients being readmitted within 30 days. This level of imbalance indicates that models trained without adjustment would likely be biased toward the majority class, resulting in poor sensitivity in identifying patients at risk of readmission. To address this, the Synthetic Minority Over-sampling Technique [26] was applied exclusively to the training set. SMOTE generates synthetic samples for the minority class by interpolating between existing cases and their nearest neighbors, thereby enriching the representation of positive cases. This approach allows the learning algorithms to better capture patterns related to readmission and improves the overall fairness of the model by reducing bias toward the majority class.

Although this method improves class balance, it may also introduce overfitting by creating artificial patterns that do not fully capture clinical variability. In practice, this can cause the model to perform well during training but less effectively in real-world applications. To address this concern, the evaluation of the model was focused on recall and F1-score using unseen data. Recall was prioritized to ensure that high-risk patients are not overlooked, while the F1-score provides a balance between precision and recall, making the evaluation more reliable in imbalanced healthcare datasets.

2.4 Model Training and Evaluation

After completing the preprocessing and class balancing stages on the training data, the process continued with model training using eight popular machine learning algorithms widely used in tabular data, namely Random Forest, Gradient Boosting, Extra Trees, XGBoost, Logistic Regression, Naïve Bayes, KNN, and Decision Tree. Each model was trained independently and adjusted through a hyperparameter tuning process using cross-validation techniques to optimize performance and prevent overfitting. After the training process, each model was evaluated on a testing set that was kept as unseen data to objectively measure its generalization ability.

To enhance predictive stability and accuracy, this study employs various ensemble learning strategies, including soft voting, hard voting, bagging, AdaBoost, stacking, and weighted ensemble. These methods integrate predictions from multiple optimized classifiers such as Random Forest, Extra Trees, Gradient Boosting, and XGBoost, trained on resampled data using SMOTE to address class imbalance. Each ensemble combines model outputs either through majority voting, probability averaging, or weighted aggregation, resulting in more balanced and robust predictions.

After obtaining the model with the optimal hyperparameters, it was evaluated on the unseen testing set using standard performance metrics, including Accuracy, Precision, Recall, and F1-Score. These metrics are computed based on the classification outcomes as defined in Formulas 1 to 4.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \quad (1)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \quad (2)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \quad (3)$$

$$\text{F1-Score} = 2\times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision}+\text{Recall})} \quad (4)$$

Formulas 1 to 4 present the formulas for Accuracy, Precision, Recall, and F1-Score. In these equations, TP refers to True Positives, TN to True Negatives, FP to False Positives, and FN to False Negatives. Precision measures the proportion of correctly predicted positive cases, Recall emphasizes sensitivity in identifying actual positive cases, and F1-Score balances both Precision and Recall.

The best-performing ensemble model is further analyzed using Explainable AI (XAI) methods, namely SHAP (SHapley Additive Explanations) [27] and LIME (Local Interpretable Model-agnostic Explanations) [28]. SHAP provides both overall feature importance and explanations for individual predictions using visual plots, while LIME explains single predictions by approximating the model locally. These techniques help reveal which features most influence the model's decision in predicting 30-day hospital readmission, improving transparency and supporting clinical interpretation.

3. Results and Discussions

3.1 Performance Comparison of Models

Based on the evaluation results shown in Table 2, all models performed competitively, with accuracy between 57.76% and 64.69%. XGBoost was the best individual model, while the weighted ensemble of the top four models achieved the highest F1-score (0.6612) and recall (0.8943), showing strong sensitivity in identifying readmissions. In medical settings, high recall is crucial to avoid missing high-risk patients, and a recent review found that 61.8% of studies on imbalanced medical data prioritize recall as a key metric[29]. Given the significant class imbalance in this study, focusing on recall is essential for accurately identifying minority class patients who are most vulnerable to readmission. Overall, ensemble methods particularly stacking and soft voting demonstrated

better metric balance, supporting their use in sensitive clinical prediction tasks.

Table 2. Model Performance Summary

Model	Metric Evaluation			
	Accuracy	Precision	Recall	F1-Score
Random Forest	0.64	0.61	0.58	0.59
Gradient Boosting	0.64	0.62	0.59	0.60
Extra Trees	0.64	0.62	0.56	0.58
XGBoost	0.65	0.62	0.60	0.61
Logistic Regression	0.62	0.62	0.47	0.54
Naïve Bayes	0.60	0.65	0.30	0.41
KNN	0.57	0.54	0.52	0.53
Decision Tree	0.61	0.58	0.52	0.55
Hard Voting	0.65	0.63	0.57	0.60
Soft Voting	0.64	0.62	0.58	0.60
Bagging	0.64	0.62	0.60	0.61
AdaBoost	0.64	0.62	0.59	0.60
Stacking	0.65	0.62	0.59	0.61
Weighted Enemble	0.58	0.55	0.89	0.66

Previous studies using the same dataset provide useful benchmarks. For instance, Gandra et al. [3], reference number 3, evaluated six machine learning models including Logistic Regression, Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost, finding CatBoost to perform best with an AUC of 0.70 and accuracy of 64.2%. Compared to these benchmarks, our framework achieves substantially higher recall of 89.43% and F1-score of 0.6612, demonstrating improved sensitivity in detecting high-risk readmissions. Additionally, unlike prior work, our approach integrates dual explainability through SHAP and LIME, providing both global and local interpretive clarity, which enhances clinical trust in the model's predictions.

3.2 Explainability Analysis

In this section, explainability analysis is conducted to understand the contribution of each feature to the model's prediction results. To achieve both global and local interpretability, we employed SHAP and LIME. SHAP provides a unified measure of feature impact across all predictions (global), while also allowing local explanations for individual cases. LIME complements this by generating interpretable approximations of the model's behavior around a specific prediction, providing detailed insights into why a particular outcome was generated.

Figure 2 shows the global feature importance based on SHAP values. The most influential features for predicting readmission were Medication Change Status and Inpatient Admissions, indicating that patients who experienced medication adjustments and had more frequent hospital stays were more likely to be readmitted. Overall, this plot suggests that medication patterns and healthcare utilization are primary drivers of readmission risk in diabetic patients.

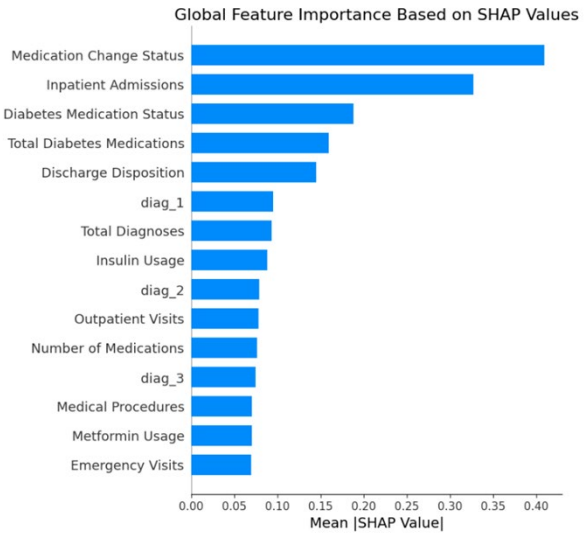


Figure 2. Global feature importance based on SHAP value

Figure 3 shows the explainability analysis using SHAP and LIME for predicting readmission and non readmission. The SHAP waterfall plot highlights that Diabetes Medication Status and Inpatient Admissions increase the likelihood of readmission, while *Medication Change Status* and *Total Diabetes Medication* reduce it. LIME results further show that Medication Change Status is the strongest factor in lowering the chance of readmission, followed by age and primary diagnosis.

LIME analysis shows that for patients predicted to be readmitted, important factors include *Medication Change Status*, *Total Diabetes Medications*, *Inpatient Admissions*, and *Tertiary Diagnosis*. These reflect changes in treatment and frequent hospital visits. For patients predicted not to be readmitted, stable treatment such as no recent medication changes and fewer inpatient visits plays a major role. For example, one patient with multiple recent admissions and medication changes was flagged as high-risk, suggesting the need for closer monitoring. In the medical context, false negative cases are particularly critical, as they represent instances where the model fails to identify patients who will be readmitted. In this study, one such case occurred despite a relatively high predicted probability of 0.449. Figure 4 presents the LIME explanation for this false negative prediction. The most influential features that positively contributed to the readmission prediction were *Medication Change Status*, followed by *Race* and *Total Diabetes Medications*, all of which pushed the decision toward the readmitted class. However, these were outweighed by strong negative contributions from *Inpatient Admissions* and *Insulin Usage*, ultimately resulting in a misclassification. This suggests that the model may have underappreciated the impact of hospitalization frequency and insulin therapy in this particular case. Although LIME successfully identified many clinically relevant features, this example highlights a key limitation of the model in balancing conflicting feature contributions, especially when the predicted probability is near the classification threshold.

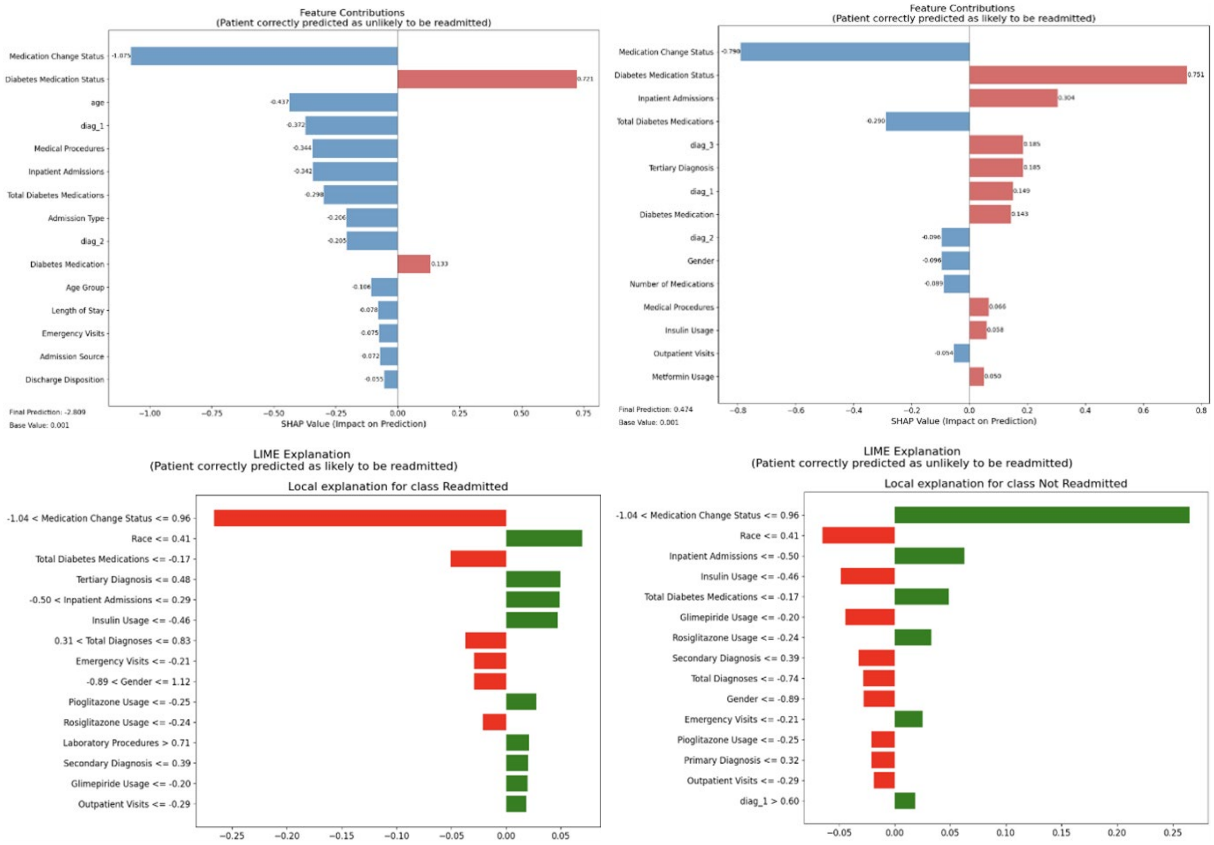


Figure 3. SHAP waterfall plot and LIME explanation for sample readmitted and not readmitted

3.3 Limitations and Future Work

The limitations of this study include: (1) reliance on the Diabetes 130 US Hospitals dataset, which may contain population-specific characteristics and institutional practices that limit generalizability to other healthcare systems; (2) potential sampling bias due to overrepresentation or underrepresentation of certain demographic or clinical profiles; (3) the use of SMOTE to address class imbalance, which may not fully capture the complexity of real-world cases and could introduce noise or risk of overfitting; and (4) the interpretability methods, SHAP and LIME, which provide valuable feature-level insights but do not reveal complex interactions, temporal dynamics, or causal relationships among variables. The proposed framework can assist hospitals in identifying high-risk diabetic patients early, enabling targeted interventions that optimize resource allocation and reduce readmission costs.

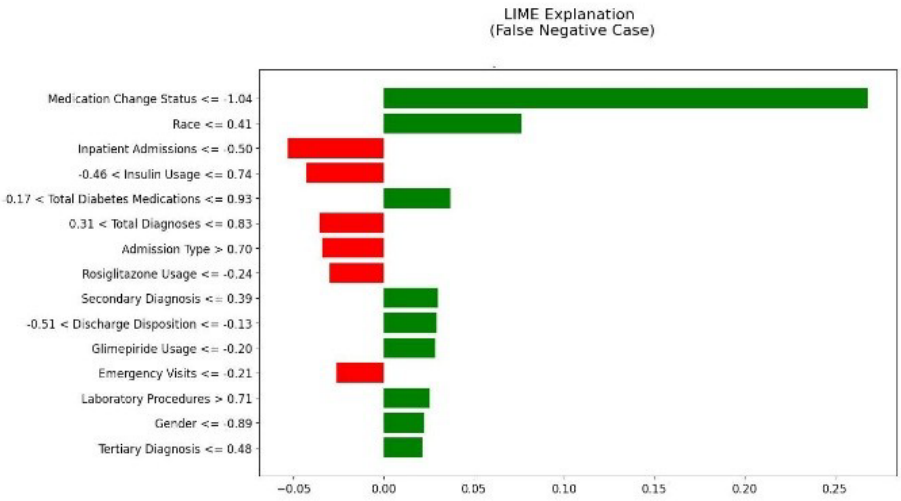


Figure 4. LIME explanation for false negative sample

To enhance clinical relevance, future research should validate the model with multi-institutional datasets, integrate it with real-time electronic health records, and include longitudinal and time-series data. Advanced interpretability methods like SHAP interaction values or counterfactual explanations should also be adopted

to improve insight and trust. Additionally, exploring real-world applications in hospitals will help support timely interventions and better resource use. Future work should focus on real-time data integration and improving clinical interpretability to enable practical adoption.

4. Conclusions

This study proposed a comprehensive machine learning framework for predicting 30-day hospital readmission in diabetic patients, combining structured data preprocessing, class balancing with SMOTE, ensemble learning strategies, and explainable AI techniques. The framework addresses key challenges in clinical prediction, including class imbalance, model performance, and interpretability. Experimental results show that while XGBoost performed well individually, the weighted ensemble outperformed other models by achieving the highest recall (0.8943) and F1-score (0.6612), making it particularly effective in identifying high-risk patients who are most vulnerable to readmission.

The use of SHAP and LIME provided valuable global and local explanations, highlighting key predictors such as *Medication Change Status* and *Inpatient Admissions* factors that are clinically intuitive and aligned with existing medical understanding. These insights not only improve model transparency but also support clinical trust in automated decision-making tools.

The contribution of this study lies in the integration of predictive performance and interpretability in a unified pipeline, demonstrating how machine learning can be effectively applied to real-world healthcare problems. The framework also reinforces the importance of handling imbalanced data and ensuring interpretability in sensitive applications like hospital readmission. By offering a solution that is both accurate and explainable, this work adds to the growing field of interpretable machine learning in healthcare. Future directions include validating the model on external datasets to test generalizability and exploring advanced interpretability methods such as SHAP interaction values. This can further improve the robustness and clinical relevance of the model, paving the way for its integration into the decision support system

References

[1] S. Chatterjee, K. Khunti, and M. J. Davies, "Type 2 diabetes," The Lancet, vol. 389, no. 10085, pp. 2239–2251, June 2017, doi: 10.1016/s0140-6736(17)30058-2.

[2] World Health Organization, "The top 10 causes of death," 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death.

[3] A. Gandra, "Predicting hospital readmissions in diabetes patients: A comparative study of machine learning models," *Int J Health Sci (Qassim)*, vol. 8, no. 3, pp. 289–297, Sep. 2024, doi: 10.53730/ijhs.v8n3.15189.

[4] O. G. Emi-Johnson and K. J. Nkrumah, "Predicting 30-Day Hospital Readmission in Patients With Diabetes Using Machine Learning on Electronic Health Record Data," *Cureus*, Apr. 2025, doi: 10.7759/cureus.82437.

[5] K. M. Dungan, "The Effect of Diabetes on Hospital Readmissions," J Diabetes Sci Technol, vol. 6, no. 5, pp. 1045–1052, Sept. 2012, doi: 10.1177/193229681200600508.

[6] M. S. Stefan *et al.*, "Hospital performance measures and 30-day readmission rates," *J Gen Intern Med*, vol. 28, no. 3, pp. 377–385, Mar. 2013, doi: 10.1007/s11606-012-2229-8.

[7] Y. Shang *et al.*, "The 30-days hospital readmission risk in diabetic patients: predictive modeling with machine learning classifiers," *BMC Med Inform Decis Mak*, vol. 21, Jul. 2021, doi: 10.1186/s12911-021-01423-y.

[8] S. Ostling *et al.*, "The relationship between diabetes mellitus and 30-day readmission rates," *Clin Diabetes Endocrinol*, vol. 3, no. 1, Dec. 2017, doi: 10.1186/s40842-016-0040-x.

[9] W. Yang *et al.*, "Economic costs of diabetes in the U.S. in 2012," *Diabetes Care*, vol. 36, no. 4, pp. 1033–1046, Apr. 2013, doi: 10.2337/dc12-2625.

[10] D. J. Rubin, "Hospital Readmission of Patients with Diabetes," Curr Diab Rep, vol. 15, no. 4, Feb. 2015, doi: 10.1007/s11892-015-0584-7.

[11] R. Miotto, L. Li, B. A. Kidd, and J. T. Dudley, "Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records," *Sci Rep*, vol. 6, May 2016, doi: 10.1038/srep26094.

[12] K. Davagdorj, J. W. Bae, V. H. Pham, N. Theera-Umpon, and K. H. Ryu, "Explainable Artificial Intelligence Based Framework for Non-Communicable Diseases Prediction," *IEEE Access*, vol. 9, pp. 123672–123688, 2021, doi: 10.1109/ACCESS.2021.3110336.

[13] I. A. Elgendy, M. Hosny, M. A. Albashrawi, and S. Alsenan, "Dual-stage explainable ensemble learning model for diabetes diagnosis," *Expert Syst Appl*, vol. 274, May 2025, doi: 10.1016/j.eswa.2025.126899.

[14] Z. Xiao, M. Wang, Y. Zhao, and H. Wang, "A Biomarker-Driven and Interpretable Machine Learning Model for Diagnosing Diabetes Mellitus," *Food Sci Nutr*, vol. 13, no. 5, May 2025, doi: 10.1002/fsn3.70234.

[15] N. G. Ramadhan, Adiwijaya, and A. Romadhony, "Preprocessing Handling to Enhance Detection of Type 2 Diabetes Mellitus based on Random Forest," IJACSA, vol. 12, no. 7, 2021, doi: 10.14569/ijacsa.2021.0120726.

[16] L. Mpofu, B. Ndlovu, S. Dube, M. Muduva, F. Jacqueline, and K. Maguraushe, "Predictive Model for Hospital Readmission of Diabetic Patients," IEOM Society International, Jul. 2024. doi: 10.46254/af05.20240252.

[17] H. Lu and S. Uddin, "Explainable Stacking-Based Model for Predicting Hospital Readmission for Diabetic Patients," *Information (Switzerland)*, vol. 13, no. 9, Sep. 2022, doi: 10.3390/info13090436.

[18] N. Alnazari *et al.*, "Development of explainable artificial intelligence based machine learning model for predicting 30-day hospital readmission after renal transplantation," *BMC Nephrol*, vol. 26, no. 1, Dec. 2025, doi: 10.1186/s12882-025-04128-w.

[19] Y. Zhang *et al.*, "Explainable machine learning for predicting 30-day readmission in acute heart failure patients," *iScience*, vol. 27, no. 7, Jul. 2024, doi: 10.1016/j.isci.2024.110281.

[20] X. Gao, S. Alam, P. Shi, F. Dexter, and N. Kong, "Interpretable machine learning models for hospital readmission prediction: a two-step extracted regression tree approach," *BMC Med Inform Decis Mak*, vol. 23, no. 1, Dec. 2023, doi: 10.1186/s12911-023-02193-5.

[21] R. Ruiz de San Martín *et al.*, "Global and Local Interpretable Machine Learning Allow Early Prediction of Unscheduled Hospital Readmission," *Mach Learn Knowl Extr*, vol. 6, no. 3, pp. 1653–1666, Jul. 2024, doi: 10.3390/make6030080.

[22] I. Tasin, T. U. Nabil, S. Islam, and R. Khan, "Diabetes prediction using machine learning and explainable AI techniques," *Healthc Technol Lett*, vol. 10, no. 1–2, pp. 1–10, Feb. 2023, doi: 10.1049/htl2.12039.

[23] B. Strack *et al.*, "Impact of HbA1c measurement on hospital readmission rates: Analysis of 70,000 clinical database patient records," *Biomed Res Int*, vol. 2014, 2014, doi: 10.1155/2014/781670.

[24] M. M. Islam, H. R. Rifat, M. S. Bin Shahid, A. Akhter, M. A. Uddin, and K. M. M. Uddin, "Explainable Machine Learning for Efficient Diabetes Prediction Using Hyperparameter

Tuning, SHAP Analysis, Partial Dependency, and LIME," *Engineering Reports*, Jan. 2024, doi: 10.1002/eng2.13080.

[25] H. Kang, "The Prevention and Handling of the Missing Data," Korean J Anesthesiol, vol. 64, no. 5, p. 402, 2013, doi: 10.4097/kjae.2013.64.5.402.

[26] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," jair, vol. 16, pp. 321–357, June 2002, doi: 10.1613/jair.953.

[27] S. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," May 2017, [Online]. Available: http://arxiv.org/abs/1705.07874/.

[28] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why Should I Trust You?' Explaining the Predictions of Any Classifier," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Association for Computing Machinery, Aug. 2016, pp. 1135–1144. doi: 10.1145/2939672.2939778.

[29] M. Salmi, D. Atif, D. Oliva, A. Abraham, and S. Ventura, "Handling imbalanced medical datasets: review of a decade of research," *Artif Intell Rev*, vol. 57, no. 10, Oct. 2024, doi: 10.1007/s10462-024-10884-2.