Received: 31 December 2024 | Revised: 23 March 2025 | Accepted: 19 May 2025 | Published online: 4 June 2025

RESEARCH ARTICLE

Medinformatics
2025, Vol. 2(3) 145–153
DOI: 10.47852/bonviewMEDIN52025119

BON VIEW PUBLISHING

A Solution Approach with Ensemble-Based Learning Technology for Predicting Early Readmission Among Patients with Heart Failure (HF) Diagnosis Using Electronic Health Records (EHR)

Muhammad Adib Uz Zaman[1,*] **and Odunayo Gabriel Adepoju**[1]

[1]*School of Information Technology, University of Cincinnati, USA*

Abstract: Heart failure is a major health issue affecting millions globally, placing a heavy burden on patients and healthcare systems. A critical challenge in managing heart failure is the high rate of early hospital readmissions. Many patients, after discharge, are readmitted within a short period, leading to further physical and emotional distress. These frequent readmissions also result in significant financial strain, both for patients and healthcare providers. Despite efforts by hospitals and government agencies, early readmission rates remain high, causing continued patient suffering and economic hardship. Health information technology (IT) provides a promising solution by utilizing data-driven approaches to predict and mitigate readmission risks. Advanced tools integrated with electronic health records (EHR) can help identify patients at higher risk of early readmission, enabling timely interventions. This approach has the potential to improve patient outcomes while alleviating the financial and logistical challenges associated with repeated hospital stays. This study explores the implementation of a Health IT solution leveraging ensemble learning, specifically an extreme gradient boost (XGBoost) algorithm, to predict early readmission risk in heart failure patients. By analyzing data from EHR, the model aims to accurately identify high-risk patients, allowing healthcare providers to take preventive measures. The findings emphasize the potential of machine learning tools to enhance healthcare efficiency and transform the management of heart failure readmissions, benefiting patients and healthcare systems alike. The XGBoost model achieved an AUC of 0.78, with a recall of 0.76 for predicting early readmissions. However, the model demonstrated high overall accuracy but struggled with lower precision (0.23) for minority class predictions due to class imbalance. The study further used SHAP to explain feature importance.

Keywords: Health IT, EHR, heart failure, XGBoost, early readmission

1. Introduction

Heart failure (HF) is a significant public health concern, contributing to high morbidity, mortality, and healthcare costs worldwide. Characterized by the heart's inability to pump sufficient blood to meet the body's metabolic needs, HF affects millions globally and is among the leading causes of hospitalizations in adults aged 65 years and older. The economic burden associated with HF is staggering, with frequent hospital readmissions accounting for a substantial portion of the costs. Early identification of patients at risk of readmission presents an opportunity to improve patient outcomes and reduce healthcare expenditure. However, accurately predicting these readmissions remains a complex challenge due to the multifactorial nature of HF and its associated comorbidities.

Hospital readmissions, particularly within 30 days of discharge, are increasingly used as a quality metric for healthcare systems. Policymakers and organizations, such as the Centers for Medicare & Medicaid Services, have implemented programs to penalize hospitals with excessive readmission rates for conditions like HF. These initiatives have spurred interest in developing effective predictive models to identify high-risk patients and implement targeted interventions. Nevertheless, the heterogeneity in patient populations, clinical presentations, and social determinants of health complicates efforts to develop accurate and generalizable predictive tools.

The significance of early readmissions in HF is particularly pronounced because of the underlying pathophysiology of HF, which is often characterized by fluctuating symptoms, recurrent exacerbations, and progressive deterioration. Inadequately managed or poorly coordinated care during the early post-discharge period may lead to worsening symptoms, which can ultimately result in the need for re-hospitalization. Early readmissions in HF patients are often associated with a higher risk of adverse outcomes, including mortality, decreased quality of life, and increased healthcare costs. Therefore, understanding the factors that contribute to early readmissions is essential for improving patient outcomes and reducing the financial burden on the healthcare system.

Recent advancements in health information technology (Health IT) and machine learning offer promising avenues for addressing this challenge. By leveraging electronic health records (EHRs), demographic data, and clinical parameters, machine learning models can analyze complex patterns and interactions that traditional statistical methods might overlook. Ensemble learning [1–4], a

*Corresponding author: Muhammad Adib Uz Zaman, School of Information Technology, University of Cincinnati, USA. Email: adib.zaman@uc.edu

machine learning paradigm that combines multiple base models to improve prediction performance, has garnered attention for its potential to enhance predictive accuracy and robustness. Ensemble methods, such as random forests [5], gradient boosting machines, and stacking, have demonstrated superior performance in various healthcare applications, including disease diagnosis, prognosis, and risk stratification.

This study explores the application of an ensemble learning-based Health IT solution to predict early readmission risk among hospitalized patients with a primary diagnosis of HF. The rationale for focusing on ensemble learning lies in its ability to mitigate overfitting, enhance model stability, and integrate diverse predictive signals from heterogeneous data sources. By constructing an ensemble model tailored to the nuances of HF readmissions, this research aims to provide a scalable and interpretable tool for clinicians and healthcare administrators.

The choice of HF as the focus of this research is motivated by its clinical significance and the unique challenges it presents. HF readmissions are influenced by a myriad of factors, including disease severity, medication adherence, social support, and healthcare access. Traditional risk prediction tools often fail to capture these complexities, leading to suboptimal performance and limited utility in clinical practice. Ensemble learning offers a promising alternative by accommodating non-linear relationships, high-dimensional data, and interactions among predictors, thereby improving the identification of high-risk patients [6].

Another critical aspect of this research is the integration of Health IT with predictive analytics. The widespread adoption of EHRs has created unprecedented opportunities to harness real-world data for predictive modeling. However, the integration of machine learning models into clinical workflows remains a significant challenge. This study aims to bridge this gap by developing a user-friendly Health IT solution that seamlessly incorporates ensemble learning predictions into the decision-making process. Such integration has the potential to facilitate timely interventions, optimize resource allocation, and improve patient outcomes.

To achieve these objectives, this research adopts a comprehensive methodology encompassing data preprocessing, feature engineering, model development, and validation. The study leverages a large, multicenter EHR dataset [7–10] containing demographic, clinical, laboratory, and social determinants of health data. Rigorous preprocessing steps ensure data quality and address common challenges, such as missing values and imbalanced classes. Feature engineering techniques are employed to extract meaningful predictors, while ensemble learning algorithms are optimized to maximize predictive performance. The models are evaluated using robust validation techniques, including cross-validation and external validation on independent datasets, to ensure generalizability and reliability.

The implications of this research extend beyond predictive modeling. By providing a scalable framework for ensemble learning in healthcare, this study contributes to the growing body of literature on machine learning applications in clinical settings. Moreover, it highlights the importance of interdisciplinary collaboration among clinicians, data scientists, and Health IT professionals to address pressing healthcare challenges. The findings of this study have the potential to inform policies and practices aimed at reducing HF readmissions, ultimately improving the quality of care and patient outcomes.

Despite the promising potential of machine learning in healthcare, several challenges warrant consideration. Ethical concerns, such as data privacy [11], algorithmic bias [12], and transparency, must be addressed to ensure the responsible use of predictive models. Additionally, the interpretability of ensemble models remains a critical factor for their adoption in clinical practice. This study prioritizes explainability by incorporating model interpretation techniques, such as SHAP (Shapley Additive Explanations) values, to elucidate the contribution of individual predictors to readmission risk. By fostering trust and understanding among clinicians, such efforts aim to bridge the gap between advanced analytics and practical implementation. In conclusion, this research investigates the development and implementation of an ensemble learning-based Health IT solution to predict early readmission risk among hospitalized HF patients. By addressing the limitations of traditional risk prediction approaches and leveraging the strengths of ensemble learning, this study aspires to advance the field of predictive analytics in healthcare. The integration of machine learning with Health IT holds great promise for transforming patient care, particularly for complex conditions like HF, where timely and accurate risk stratification can make a profound difference. This introduction sets the stage for the subsequent sections of the article, which detail the methodology, results, and implications of this innovative approach. Besides, the study also mentions several important precision improvement issues that occur in an imbalanced dataset (which translates into rare event prediction in most cases). This is a special type of problem that is very difficult to deal with traditional machine learning or data analysis approaches.

This research study aims to make the following unique contributions:

1) Prediction of early readmissions: The model aims to identify high-risk patients after discharge, enabling timely interventions to reduce hospital readmissions, improving patient outcomes, and alleviating healthcare system burdens.
2) Machine learning for Health IT: The research highlights the potential of advanced machine learning tools integrated with Health IT to transform HF management and reduce the financial strain caused by repeated hospital stays.
3) Use of SHAP for feature explanation: The study employs SHAP to interpret and explain the feature importance in the predictive model, enhancing transparency and interpretability of machine learning predictions.

This research article has been segmented into different sections for easier readability and organization. Section 1 combines the introduction of HF early readmission and related background information. Section 2 discusses the materials and methods that are used in this study followed by Section 3 which details the numerical results and discussions from the proposed model. Section 4 includes the conclusions and future directions of this potentially valuable research area.

2. Materials and Methods

The dataset was organized into five major categories: laboratory results, medications, outcomes, baseline clinical features and comorbidities, and demographic details. When patients were admitted to the hospital for the first time, their demographic information was manually entered into the electronic medical record (EMR) system by nurses. For returning patients, this information was automatically retrieved from prior records, though nurses carefully reviewed and corrected any missing or inaccurate details they identified [9]. To improve accuracy and consistency during data entry, dropdown menus were integrated into the EMR system for certain variables, such as sex, admission department, and occupation. Physicians and laboratory personnel electronically recorded medication details and test results. SQL queries were then used to extract the data from the EMR system to create the database, with manual validation performed by

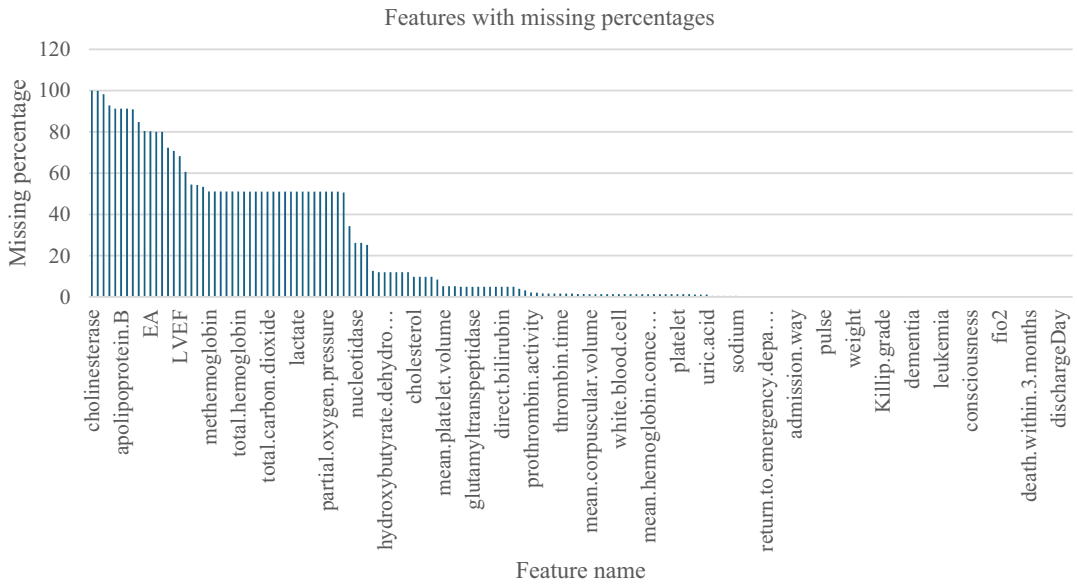

Figure 1. Original features from the dataset with their missing percentages

randomly reviewing 50 patient records for accuracy. A notable challenge arose due to language barriers, as much of the data, including drug names, diagnoses, and lab test results, were recorded in Chinese within the EMR system. There was a huge number of missing values across many columns as shown in Figure 1.

Demographic information extracted from the EMR included age, sex, height, weight, marital status, occupation, admission ward, admission type (emergency or non-emergency), discharge department, admission date, and the number of previous hospital visits. These details were captured from the first sheet of each patient's medical records [9].

Baseline clinical data documented on the day of admission included vital signs such as body temperature, pulse, respiration rate, systolic and diastolic blood pressure, and mean arterial pressure. Additional measurements included weight, height, body mass index, and the Glasgow Coma Scale score. Clinical indicators such as the type of HF, New York Heart Association functional classification, and Killip grade (ranging from Class 1, indicating no rales or third heart sound, to Class 4, indicating cardiogenic shock) were also recorded. Advanced assessments, such as echocardiography, provided metrics like tricuspid regurgitation velocity, pressure readings, mitral valve E and A wave velocities, E/A ratio, left ventricular ejection fraction, and left ventricular end-diastolic diameter [9].

Comorbidities [13, 14] documented in the admission notes included myocardial infarction, congestive HF, peripheral vascular disease, cerebrovascular disease, dementia, chronic obstructive pulmonary disease, connective tissue disorders, peptic ulcer disease, diabetes, chronic kidney disease (moderate to severe), hemiplegia, leukemia, malignant lymphoma, solid tumors, liver disease, and AIDS. These diagnoses were used to compute the Charlson Comorbidity Index. Furthermore, some patients received new diagnoses during their hospital stay, including cases of congestive HF that were not present before admission.

Laboratory results from the first day of admission provided a wide range of clinical data. Cardiac biomarkers such as brain natriuretic peptide and high-sensitivity troponin were measured, along with coagulation profiles that included tests like D-dimer, INR, prothrombin time, and thrombin time. Electrolyte levels for calcium, potassium, chloride, sodium, and magnesium were recorded, alongside enzyme levels such as creatine kinase, lactate dehydrogenase, and transaminases. Markers of renal function, including serum creatinine, urea, uric acid, glomerular filtration rate, and cystatin, were also evaluated. Other laboratory measurements encompassed lipid profiles, protein levels, bilirubin, blood gas analysis, and metabolic markers such as lactate and glucose. These results offered valuable insights into the clinical condition of the patients at the time of admission [9].

Figure 2 presents the correlation of various features with early readmission, with the y-axis representing correlation scores (ranging approximately from −0.15 to 0.15) and the x-axis listing all features. Overall, the correlations are relatively weak, with most values hovering near zero. This suggests that no individual feature has a strong linear relationship with early readmission, and the outcome is likely influenced by a combination of multiple factors rather than a single dominant variable.

Both positive and negative correlations are observed, albeit with low magnitudes. Features with positive correlations indicate that higher values might slightly increase the likelihood of early readmission, while negative correlations suggest the opposite. However, the weak strength of these relationships means that these features, on their own, provide limited predictive power. The variability in correlation scores across features indicates that some variables contribute more meaningfully than others, though none stand out as particularly influential.

These findings have important implications for modeling. The weak correlations imply that linear models may struggle to capture the complexity of early readmission, as relationships between features and the target variable may be more nuanced or non-linear. More sophisticated approaches, such as decision trees, random forests [15], or neural networks, could better account for feature interactions and non-linear dependencies. Additionally, while features with higher positive or negative correlations may hold some predictive value, those with near-zero correlations may add noise and should be considered for removal during feature selection.

To improve performance, techniques like recursive feature elimination have been used to identify irrelevant features and streamline the dataset. At the same time, handling class imbalance

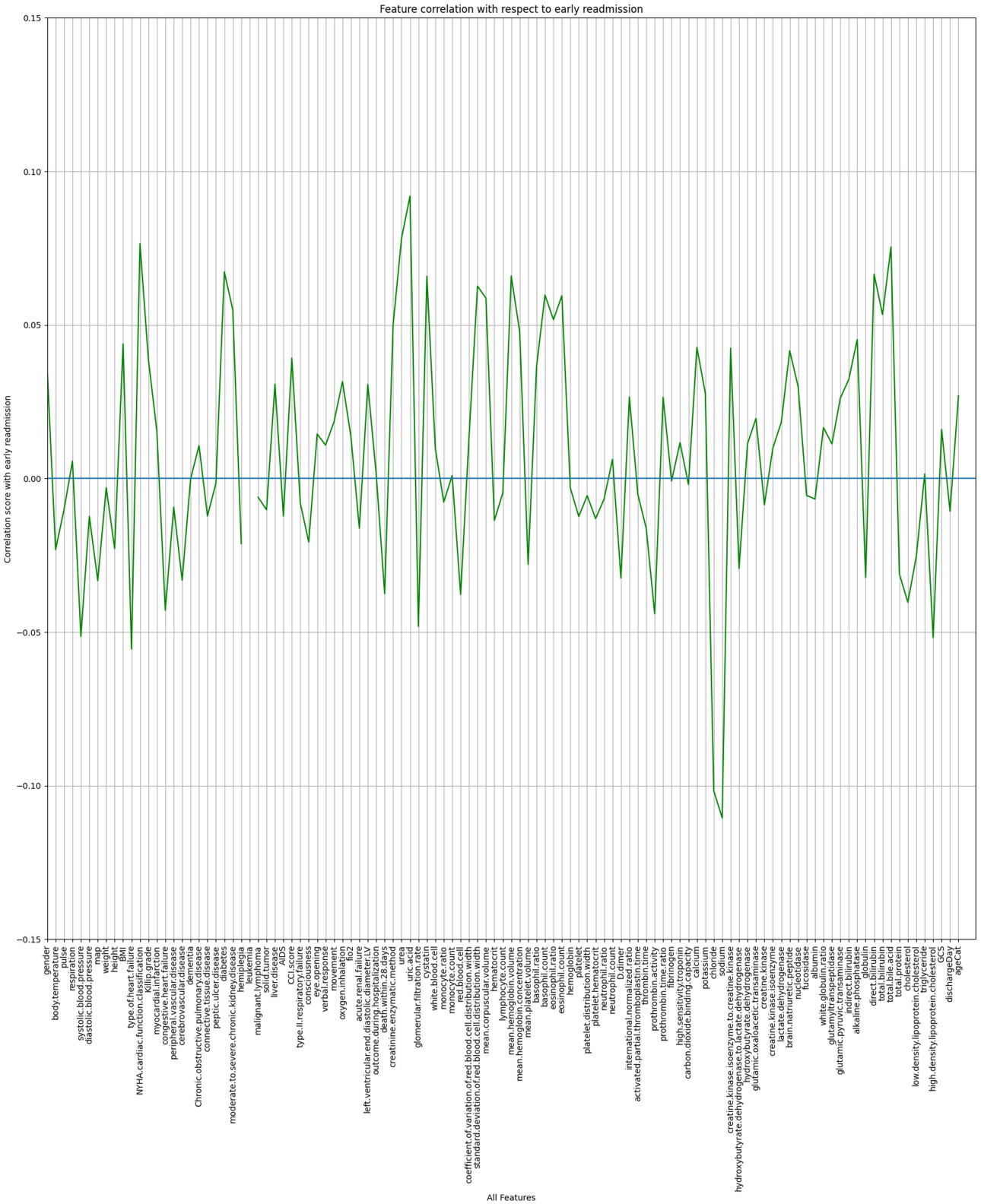

Figure 2. Numerical feature correlation with respect to early readmission

might shift the importance of certain features and impact model outcomes. In summary, the weak linear relationships evident in the figure suggest that early readmission prediction requires not only advanced modeling techniques but also careful consideration of feature selection and preprocessing strategies. The following sub-sections discuss the methods that have been used in this study.

2.1. Feature selection

Recursive feature elimination (RFE) is a feature selection technique used to improve the performance and interpretability of machine learning models. It works by recursively removing less important features from the dataset, based on their contribution to

the predictive model, until the optimal subset of features is identified. RFE is particularly effective in high-dimensional datasets where irrelevant or redundant features may negatively impact model performance.

The process begins by fitting a machine learning model, such as a linear model or a tree-based model, to the entire set of features. The importance of each feature is evaluated using the model's internal metrics, such as coefficients or feature importance scores. The least significant feature, as determined by the evaluation metric, is removed from the dataset. This process is repeated iteratively, with the model being refitted at each step, until the desired number of features is reached.

RFE helps reduce overfitting, improves computational efficiency, and can make the model more interpretable by retaining only the most relevant features. It is widely used in applications such as biomedical research, finance, and marketing, where understanding the impact of key variables is essential. However, RFE's computational expense may increase for large datasets, especially when paired with complex models. In this study, we selected 17 features using RFE for predictive modeling.

2.2. Synthetic minority oversampling technique (SMOTE)

SMOTE [16, 17] is a popular method used to address the issue of class imbalance in datasets, a common challenge in machine learning tasks. Class imbalance occurs when one class (minority class) has significantly fewer samples compared to the other class (majority class). Such imbalances can lead to biased model predictions, as the algorithm often prioritizes the majority class.

SMOTE tackles this problem by oversampling the minority class using synthetic data generation. Unlike simple oversampling, which duplicates existing minority samples, SMOTE generates new synthetic samples by interpolating between existing instances of the minority class. The process involves selecting a random minority class instance and one of its k-nearest neighbors, creating a synthetic sample along the line connecting these two points in the feature space.

2.3. Ensemble-based learning

The ensemble-based learning that was selected for this study is XGBoost that is short for eXtreme Gradient Boosting [18]. It is a powerful machine learning algorithm based on gradient boosting, designed to improve both computational speed and predictive performance. It has been widely used for classification tasks due to its efficiency and scalability. XGBoost builds an ensemble of decision trees sequentially, optimizing an objective function through gradient descent. Below, we provide brief details into the mathematical details underlying the XGBoost classifier [18].

The following equation shows the objective function (loss) to be minimized by the XGBoost classifier [18].

$$\mathcal{L}(\theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{1}$$

Where:

$l(y_i, \hat{y}_i)$ is the loss function measuring the difference between the predicted value and the true label

$\Omega(f_k)$ is the regularization term that controls the complexity of the model

θ represents the model parameters.

n is the number of training samples, and

K is the total number of trees.

XGBoost constructs trees iteratively by adding one weak learner at a time to minimize the objective function. The prediction at the t^{th} iteration is given by:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i) \tag{2}$$

The regularization term $\Omega(f_k)$ controls the complexity of the tree and helps prevent overfitting. For a tree with T leaf nodes, it is defined as:

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \tag{3}$$

Where:

γ is the penalty for the number of leaf nodes T

λ is the L^2 regularization term

To grow a decision tree, XGBoost evaluates candidate splits by calculating the gain in the objective function [18]. The gain for a split is computed as:

$$\text{Gain} = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \gamma \tag{4}$$

Where:

G_L and G_R are the sums of gradients for the left and right child nodes, respectively

H and H_R are the sums of Hessians for the left and right child nodes, respectively.

The split with the highest gain is selected, and the process continues recursively until a stopping criterion (e.g., maximum depth or minimum leaf size) is met.

XGBoost also introduces a shrinkage parameter η (learning rate) to scale the contribution of each tree [18]. After fitting a tree, the model updates the predictions as:

$$\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + \eta f_t(x_i) \tag{5}$$

where $0 < \eta \leq 1$. Smaller values of η often lead to better generalization but require more iterations.

3. Results and Discussion

The prediction process followed ten times ten-fold approach for cross-validation. Figure 3 shows the confusion matrix that visualizes the performance of an XGBoost classification model using 10 selected features on unseen test data set which was 10% of the total data. It provides an overview of the predictions made by the model compared to the actual target labels, helping to assess its accuracy and error patterns. The rows represent the predicted classes, while the columns denote the actual classes. The classes are labeled as "0" (no early readmission) and "1" (early readmission).

True positives (TP): The bottom-right cell indicates 13 cases where the model correctly predicted the early readmission class (1). These are instances where the model's prediction aligned with the actual early readmission class, demonstrating its capability to identify TP effectively.

True negatives (TN): The top-left cell shows 136 cases where the model correctly identified the no early readmission class (0). This indicates the model's reliability in predicting negative outcomes accurately.

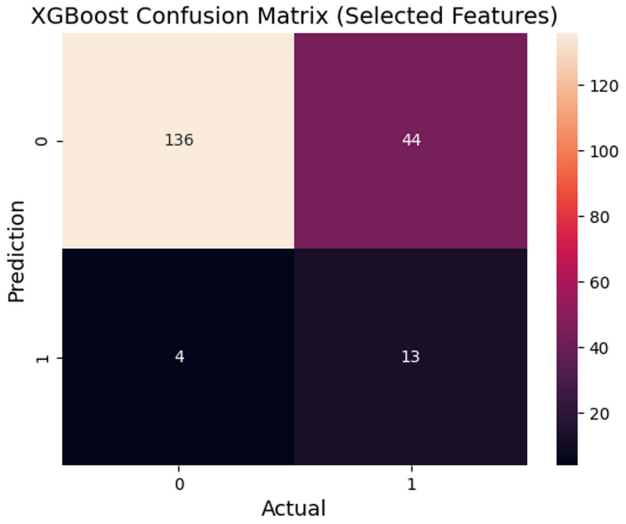

Figure 3. Confusion matrix for testing

False positives (FP): The top-right cell contains 44 instances where the model incorrectly classified non-early readmission cases (actual 0) as early readmission cases (1). These errors reflect the model's tendency to misclassify some negatives.

False negatives (FN): The bottom-left cell indicates 4 cases where the model failed to identify the early readmission class (1) and incorrectly predicted them as no early readmitted class (0).

This confusion matrix [19, 20] highlights the strengths of the XGBoost model, particularly its reasonable recall and accuracy. However, the relatively lower precision suggests potential improvements, especially in reducing FNs. This could involve techniques such as hyperparameter tuning, or more enhanced feature engineering to improve the model's sensitivity to positive cases.

The following metric can be derived from these values like the following:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{6}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{8}$$

$$\text{F1} - \text{Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{9}$$

The classification report for testing is shown below:

The classification report in Table 1 evaluates the performance of a model predicting early readmission in patients. It measures the model's performance using precision, recall, and *F*1 score for each class: "No early readmission" and "Early readmission." Additionally, macro and weighted averages summarize the metrics across all classes.

Table 1. Classification report for XGBoost classifier

Class	Precision	Recall	*F*1 score	Support
No early readmission	0.97	0.76	0.85	180
Early readmission	0.23	0.76	0.35	17
Macro average	0.60	0.76	0.60	197
Weighted average	0.91	0.76	0.81	197

No Early Readmission class has a high precision (0.97), indicating the model effectively identifies TNs with minimal FPs. However, the recall (0.76) suggests that 24% of the actual cases are missed. The *F*1 score of 0.85 balances these metrics, showing reasonable performance for this majority class.

Early readmission (minority) class has a low precision (0.23), meaning a significant proportion of predicted positive cases are FPs. However, recall is relatively high (0.76), demonstrating the model captures most actual early readmissions. The *F*1 score (0.35) highlights the model's difficulty in balancing precision and recall, likely due to the class imbalance.

When calculating performance metrics like precision, recall, and *F*1 score, the macro-average approach treats both classes equally, regardless of how many samples each class has. This means that the metrics are averaged across all classes, with each class contributing the same weight to the final result. The scores (0.60 for precision and *F*1, 0.76 for recall) indicate mediocre model performance for the minority class. Also, the weighted average adjusts the metrics based on class support, favoring "No early readmission." The weighted *F*1 score (0.81) reflects stronger performance for the majority class while downplaying the minority class issues.

The model's performance is skewed towards the majority class. Although this imbalance was addressed using sampling, weighted loss functions, it could not improve predictions for the minority class that much.

Figure 4 represents the receiver operating characteristic (ROC) curve [21, 22], a key performance evaluation metric used in binary classification problems. The ROC curve plots the true positive rate (TPR, also called sensitivity or recall) on the y-axis against the false positive rate (FPR) on the x-axis. The FPR is defined as the proportion of negative instances incorrectly classified as positive, and the TPR quantifies the model's ability to correctly identify positive cases.

This ROC curve and the associated AUC of 0.78 demonstrate the high effectiveness of the evaluated classifier. The curve's shape and the high AUC value indicate that the model is capable of distinguishing between positive and negative classes with minimal misclassification. This performance evaluation is a strong indication that the classifier is suited for the task it is applied to, assuming the dataset and evaluation process are representative of the real-world use case. Future steps may include fine-tuning thresholds or comparing this model to alternatives to optimize performance further. However, models might not provide the best estimation if the dataset is imbalanced and only ROC-AUC curve shows good performance [23].

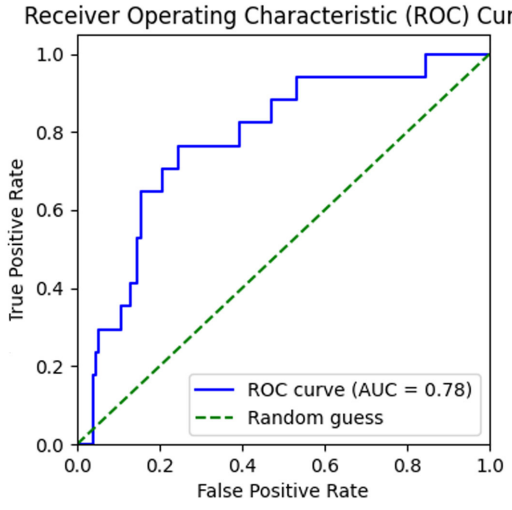

Figure 4. ROC-AUC

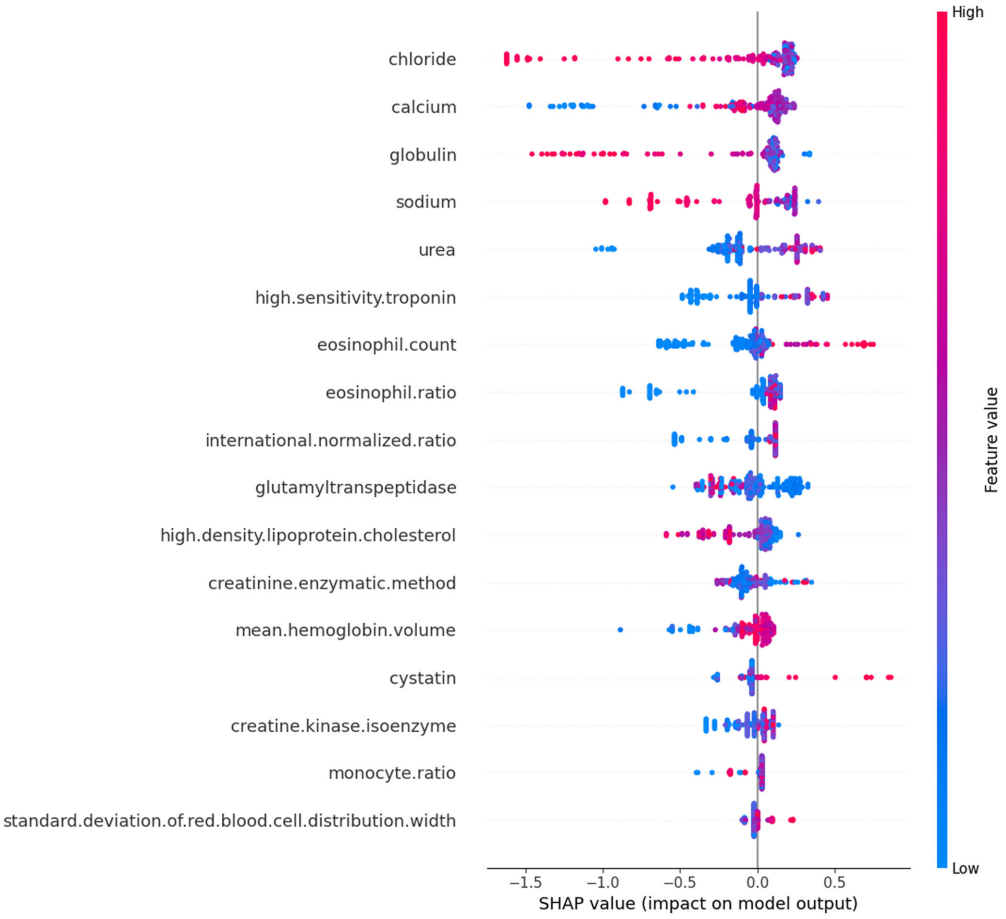

Figure 5. Feature importance by SHAP values

Figure 5 provides a detailed visualization of the significance of various features in a machine learning model using SHAP values [24]. SHAP values are a method of interpreting the predictions made by machine learning models, revealing the contribution of each feature to the final prediction. The plot is divided into several elements that contribute to its comprehensive scientific interpretation. The image is basically a bar plot, with the x-axis representing the mean SHAP values and the y-axis listing the various features. The mean SHAP values on the x-axis range from −1.5 to 0.5, denoting the average impact of each feature on the model's output. Positive SHAP values indicate a positive impact on the prediction, while negative values suggest a negative influence.

Each feature in the plot is represented by a horizontal line or a cluster of points spread along the x-axis. The density and spread of these points illustrate the distribution of SHAP values for each feature. Features with a wider spread of SHAP values have a more significant impact on the model's output. A color gradient from blue to red overlays the plot, adding another layer of information. Blue represents low feature values, while red indicates high feature values. This gradient allows us to discern how different levels of each feature influence the model's predictions. For instance, high values of cystatin (red points) have a positive impact on the model's output, while low values (blue points) have a negative impact. The plot indicates that the cystatin feature has a wide spread of SHAP values, signaling its substantial impact on the model's predictions. Eosinophil also shows a significant spread, suggesting it is another crucial feature. On the other hand, features like monocyte ratio, standard deviation of red blood cell distribution width, etc. show a smaller impact, as indicated by their narrower spread of SHAP values.

4. Conclusion and Future Directions

Health IT solutions have great potential for predicting and reducing early readmission risks in HF patients. EHRs, predictive analytics, and telemonitoring are currently the most used approaches, with evidence backing their effectiveness. However, challenges like data integration, privacy, and variability in healthcare settings need to be tackled to fully harness these technologies' potential. Innovations like AI, personalized medicine, blockchain, mHealth, NLP, and community-based interventions show promise for the future [25]. By using these technologies, healthcare providers can develop more accurate predictive models, improve patient outcomes, and lessen the burden of HF on healthcare systems. Bringing Health IT solutions into everyday clinical practice requires teamwork among healthcare providers, tech developers, policymakers, and patients. Continuous evaluation and improvement of these solutions are crucial to ensure their effectiveness and sustainability. Moving forward, the goal should be a patient-centered healthcare system that uses IT to deliver high-quality, personalized care. Focusing on early identification and intervention, Health IT solutions can play a key role in improving the lives of HF patients and addressing one of modern healthcare's most pressing challenges. The future of HF early readmission management looks bright, with

technology leading the way for more efficient, effective, and patient-centered care.

Besides, incorporating social determinants of health (SDOH) and patient-reported outcomes (PROs) into predictive models for HF readmissions can significantly enhance prediction accuracy and provide a more holistic understanding of the factors that influence patient outcomes. SDOH, including socioeconomic status, access to healthcare, housing stability, education, and social support, can profoundly impact a patient's ability to manage their condition and follow treatment plans effectively. For instance, patients with limited financial resources may struggle to afford medications or attend follow-up appointments, increasing their likelihood of readmission. Similarly, those with poor social support may face challenges in adhering to lifestyle modifications or managing stress, which can exacerbate their HF symptoms. Patient-reported outcomes, such as self-reported symptoms, quality of life, mental health status, and functional limitations, offer direct insights into the patient's perspective on their health, which often correlates with clinical outcomes. These factors are typically not captured in traditional clinical data but are crucial for understanding a patient's risk for early readmission. By integrating SDOH and PROs into predictive models, healthcare providers can develop more personalized, targeted interventions that not only address medical needs but also the broader social and emotional factors that affect health outcomes. This approach moves beyond the clinical setting to better predict and mitigate the risks of readmission, ensuring a more comprehensive and accurate assessment of patient risk.

Acknowledgement

The authors are grateful to Zigong Fourth People's Hospital for compiling the dataset (https://physionet.org/content/heart-failure-zigong/1.3/).

AI Tool Usage Statement

The authors have used AI tool for text formatting, proofreading, debugging, and refinement only. The ideas, experiments, figures, tables, etc., are not generated with AI. All these items are thoroughly checked by the authors to the best of their abilities.

Funding Support

This work is partially supported by the new faculty start-up grant from the dean of the College of Criminal Justice, Education, and Human Services (CECH) at the University of Cincinnati (UC).

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are available (upon request) in Physionet website at Hospitalized patients with heart failure: integrating electronic healthcare records and external outcome data. The GitHub code repository link for the implementation can be found at https://github.com/adibML007/XGBoost_early_readmission_paper.

Author Contribution Statement

Muhammad Adib Uz Zaman: Conceptualization, Methodology, Software, Validation, Resources, Writing – original draft, Writing – review & editing, Supervision, Project administration, Funding acquisition. **Odunayo Gabriel Adepoju:** Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Visualization.

References

[1] Namamula, L. R., & Chaytor, D. (2024). Effective ensemble learning approach for large-scale medical data analytics. *International Journal of System Assurance Engineering and Management*, *15*(1), 13–20. https://doi.org/10.1007/s13198-021-01552-7

[2] Rane, N., Choudhary, S. P., & Rane, J. (2024). Ensemble deep learning and machine learning: Applications, opportunities, challenges, and future directions. *Studies in Medical and Health Sciences*, *1*(2), 18–41. https://doi.org/10.48185/smhs.v1i2.1225

[3] Ali, F., El-Sappagh, S., Islam, S. M. R., Kwak, D., Ali, A., Imran, M., & Kwak, K. S. (2020). A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion*, *63*, 208–222. https://doi.org/10.1016/j.inffus.2020.06.008

[4] Kumar, R., Maheshwari, S., Sharma, A., Linda, S., Kumar, S., & Chatterjee, I. (2024). Ensemble learning-based early detection of influenza disease. *Multimedia Tools and Applications*, *83*(2), 5723–5743. https://doi.org/10.1007/s11042-023-15848-2

[5] Dorado-Guerra, D. Y., Corzo-Pérez, G., Paredes-Arquiola, J., & Pérez-Martín, M. Á. (2023). Machine learning models to predict nitrate concentration in a river basin. *Environmental Research Communications*, *4*(12), 125012. https://iopscience.iop.org/article/10.1088/2515-7620/acabb7

[6] Ihnaini, B., Khan, M. A., Abbas Khan, T., Abbas, S., Daoud, M. S., Ahmad, M., & Khan, M. A. (2021). A smart healthcare recommendation system for multidisciplinary diabetes patients with data fusion based on deep ensemble learning. *Computational Intelligence and Neuroscience*, *2021*(1), 4243700. https://doi.org/10.1155/2021/4243700

[7] Bardhan, I., Oh, J. H., Zheng, Z., & Kirksey, K. (2015). Predictive analytics for readmission of patients with congestive heart failure. *Information Systems Research*, *26*(1), 19–39. https://doi.org/10.1287/isre.2014.0553

[8] Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., . . . , & Eugene Stanley, H. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, *101*(23), e215–e220. https://doi.org/10.1161/01.CIR.101.23.e215

[9] Zhang, Z., Cao, L., Chen, R., Zhao, Y., Lv, L., Xu, Z., & Xu, P. (2021). Electronic healthcare records and external outcome data for hospitalized patients with heart failure. *Scientific Data*, *8*(1), 46. https://doi.org/10.1038/s41597-021-00835-9

[10] Zaman, M. A. U. (2024). Assessing different diagnoses in MIMIC-IV v2.2 and MIMIC-IV-ED datasets. *Archives of Proteomics and Bioinformatics*, *4*(1), 1–5. https://doi.org/10.33696/Proteomics.4.014

[11] Manikandan, A., Sanjay, T., Menon, G., Aswin, R., Bhaskar, P. B., Mahadev Govind, R., & Ramprasad, O. G. (2025). Issues and challenges in security and privacy with e-Healthcare: A thorough literature analysis. In N. Goel &

R. K. Yadav (Eds.), *Internet of Things enabled machine learning for biomedical applications* (pp. 222–247). CRC Press. https://doi.org/10.1201/9781003487647-13

[12] Jones, C., Castro, D. C., de Sousa Ribeiro, F., Oktay, O., McCradden, M., & Glocker, B. (2024). A causal perspective on dataset bias in machine learning for medical imaging. *Nature Machine Intelligence*, *6*(2), 138–146. https://doi.org/10.1038/s42256-024-00797-8

[13] Gomez-Ochoa, S. A., Lanzer, J. D., & Levinson, R. T. (2025). Disease network-based approaches to study comorbidity in heart failure: Current state and future perspectives. *Current Heart Failure Reports*, *22*(1), 6. https://doi.org/10.1007/s11897-024-00693-7

[14] Frank, H. A., & Karim, M. E. (2025). Physical comorbidity is associated with overnight hospitalization in US adults with asthma: An assessment of the 2005–2018 National Health and Nutrition Examination Surveys. *Journal of Asthma*, *62*(1), 155–166. https://doi.org/10.1080/02770903.2024.2393677

[15] Pham, Q. B., Tran, D. A., Ha, N. T., Towfiqul Islam, A. R. M., & Salam, R. (2022). Random forest and nature-inspired algorithms for mapping groundwater nitrate concentration in a coastal multi-layer aquifer system. *Journal of Cleaner Production*, *343*, 130900. https://doi.org/10.1016/j.jclepro.2022.130900

[16] Elreedy, D., Atiya, A. F., & Kamalov, F. (2024). A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Machine Learning*, *113*(7), 4903–4923. https://doi.org/10.1007/s10994-022-06296-4

[17] Liaw, L. C. M., Tan, S. C., Goh, P. Y., & Lim, C. P. (2025). A histogram SMOTE-based sampling algorithm with incremental learning for imbalanced data classification. *Information Sciences*, *686*, 121193. https://doi.org/10.1016/j.ins.2024.121193

[18] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

[19] Cadey, F. (2024). Machine learning classification. In F. Cady (Ed.), *The data science handbook* (2nd ed., pp. 89–107). Wiley.

[20] Arpita, H. D., Al Ryan, A., Hossain, M. F., Rahman, M. S., Sajjad, M., & Islam Prova, N. N. (2025). Exploring Bengali speech for gender classification: Machine learning and deep learning approaches. *Bulletin of Electrical Engineering and Informatics*, *14*(1), 328–337. https://doi.org/10.11591/eei.v14i1.8146

[21] Roumeliotis, S., Schurgers, J., Tsalikakis, D. G., D'Arrigo, G., Gori, M., Pitino, A., . . . , & Liakopoulos, V. (2024). ROC curve analysis: A useful statistic multi-tool in the research of nephrology. *International Urology and Nephrology*, *56*(8), 2651–2658. https://doi.org/10.1007/s11255-024-04022-8

[22] Ozel, C. B., Yakut Ozdemir, H., Dural, M., Al, A., Yalvac, H. E., Mert, G. O., . . . , & Cavusoglu, Y. (2025). The one-minute sit-to-stand test is an alternative to the 6-minute walk test in patients with atrial fibrillation: A cross-sectional study and ROC curve analysis. *International Journal of Cardiology*, *419*, 132713. https://doi.org/10.1016/j.ijcard.2024.132713

[23] Liu, Y., Li, Y., & Xie, D. (2024). Implications of imbalanced datasets for empirical ROC-AUC estimation in binary classification tasks. *Journal of Statistical Computation and Simulation*, *94*(1), 183–203. https://doi.org/10.1080/00949655.2023.2238235

[24] Alam, S. M. K., Li, P., Rahman, M., Fida, M., & Elumalai, V. (2025). Key factors affecting groundwater nitrate levels in the Yinchuan Region, Northwest China: Research using the eXtreme Gradient Boosting (XGBoost) model with the SHapley Additive exPlanations (SHAP) method. *Environmental Pollution*, *364*, 125336. https://doi.org/10.1016/j.envpol.2024.125336

[25] Nguyen, D. K., Lan, C. H., & Chan, C. L. (2021). Deep ensemble learning approaches in healthcare to enhance the prediction and diagnosing performance: The workflows, deployments, and surveys on the statistical, image-based, and sequential datasets. *International Journal of Environmental Research and Public Health*, *18*(20), 10811. https://doi.org/10.3390/ijerph182010811

How to Cite: Zaman, M. A. U., & Adepoju, O. G. (2025). A Solution Approach with Ensemble-Based Learning Technology for Predicting Early Readmission Among Patients with Heart Failure (HF) Diagnosis Using Electronic Health Records. *Medinformatics*, *2*(3), 145–153. https://doi.org/10.47852/bonviewMEDIN52025119