Received 28 January 2025, accepted 5 February 2025, date of publication 11 February 2025, date of current version 19 February 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3541069

RESEARCH ARTICLE

Predicting the Classification of Heart Failure Patients Using Optimized Machine Learning Algorithms

MARZIA AHMED[1,2], MOHD HERWAN SULAIMAN[2], (Senior Member, IEEE), MD MARUF HASSAN[1,3], (Member, IEEE), AND TOUHID BHUIYAN[4]

[1]Department of Software Engineering, Daffodil International University (DIU), Daffodil Smart City, Dhaka 1216, Bangladesh
[2]Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), Pekan 26600, Malaysia
[3]Department of Computer Science and Engineering, Southeast University, Dhaka, Bangladesh
[4]School of IT, Washington University of Science and Technology (WUST), Alexandria, VA 22314, USA

Corresponding author: Marzia Ahmed (ahmed.marzia32@gmail.com)

This work was supported by the School of IT, Washington University of Science and Technology (WUST), in Virginia, USA, Program for Scientific Publication.

ABSTRACT Heart failure is a critical condition with a high mortality rate, making accurate survival prediction essential for timely interventions. This study proposes an optimized machine learning approach using Gradient Boosting Machine (GBM) and Adaptive Inertia Weight Particle Swarm Optimization (AIW-PSO) to predict heart failure survival. The dataset, sourced from Kaggle, includes clinical features such as age, ejection fraction, and serum creatinine levels for 299 heart failure patients. To address the imbalance in survival outcomes, Synthetic Minority Over-sampling Technique (SMOTE) was employed to balance the dataset, followed by SelectKBest and Chi-square feature selection methods to retain the most significant predictors. The optimized hyperparameters for the GBM model were identified using the AIW-PSO algorithm, which effectively balanced exploration and exploitation by adaptively adjusting inertia weights. Model selection was further refined using information criteria, including Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), ensuring that the best-performing model was chosen based on both predictive accuracy and model complexity. The optimized GBM model achieved a test accuracy of 94%, demonstrating superior performance compared to traditional machine learning models. The study underscores the importance of hyperparameter tuning through metaheuristic algorithms and highlights the potential of AIW-PSO in enhancing model performance for clinical prediction tasks. These findings have significant implications for clinical decision-making, offering a reliable and interpretable tool for predicting patient outcomes in heart failure management.

INDEX TERMS Heart failure survival prediction, machine learning algorithms, hyperparameter optimization, class imbalance handling, AIW-PSO optimization.

ABBREVIATIONS AND TERMS

- **AIWPSO**: Adaptive Inertia-Weight Particle Swarm Optimization
- **GBM**: Gradient Boosting Machine
- **SMOTE**: Synthetic Minority Over-sampling Technique
- **ML**: Machine Learning
- **AUC**: Area Under the Curve
- **ROC**: Receiver Operating Characteristic
- **SVM**: Support Vector Machine
- **HF**: Heart failure
- **GA**: Genetic Algorithm
- **ADASYN**: Adaptive Synthetic
- **ANOVA**: Analysis of Variance
- **AIC**: Akaike Information Criterion
- **SBIC**: Schwarz Bayesian Information Criterion
- **HQIC**: Hannan-Quinn Criterion

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

I. INTRODUCTION

Cardiovascular diseases (CVDs) include a range of disorders affecting the heart and blood vessels, such as coronary heart disease, stroke, and heart failure (HF). The World Health

Organization (WHO) reports that CVDs are the leading cause of death globally, accounting for approximately 17.9 million deaths each year, which represents nearly 32% of all deaths worldwide.

Heart failure (HF) remains a formidable challenge in cardiovascular medicine, affecting an estimated 64.3 million people worldwide and accounting for a substantial portion of global healthcare expenditure [1]. This chronic, progressive condition is characterized by the heart's inability to pump blood efficiently, leading to a cascade of symptoms that significantly impair quality of life and elevate mortality risk. It is frequently caused by underlying conditions like diabetes, hypertension, or other heart diseases [2]. Despite advancements in therapeutic interventions, the 5-year mortality rate for HF patients hovers around 50%, underscoring the urgent need for improved prognostic tools [3].

In recent years, the intersection of machine learning (ML) and clinical medicine has opened new avenues for enhancing patient care through data-driven decision support systems [4], [34]. The application of ML algorithms to predict HF outcomes has shown promise, yet challenges persist in model accuracy and generalizability [5]. A critical bottleneck in leveraging ML for clinical prediction tasks lies in the optimization of model hyperparameters, a process that can significantly influence predictive performance [6].

The advent of metaheuristic optimization algorithms, such as Particle Swarm Optimization (PSO), has provided a powerful framework for navigating the complex hyperparameter landscape of ML models [7]. However, the classical PSO algorithm often struggles with the delicate balance between exploration and exploitation, potentially leading to suboptimal solutions [8]. To address this limitation, we propose the application of Adaptive Inertia Weight Particle Swarm Optimization (AIW-PSO), an enhanced variant that dynamically adjusts its search behavior, to optimize the hyperparameters of a Gradient Boosting Machine (GBM) model for HF survival prediction.

Our study leverages a curated dataset of 299 HF patients, encompassing a rich tapestry of clinical features including left ventricular ejection fraction, serum creatinine levels, and comorbidities [9]. To mitigate the inherent class imbalance typical of survival data, we employ the Synthetic Minority Over-sampling Technique (SMOTE), ensuring a balanced representation of outcomes [10]. Feature selection is performed using the SelectKBest algorithm in conjunction with Chi-square statistical tests, distilling the most salient predictors from the feature space [11].

The novelty of our approach lies in the synergistic integration of AIW-PSO with GBM, a powerful ensemble learning method known for its robustness in handling complex, non-linear relationships [12]. By harnessing the adaptive capabilities of AIW-PSO, we aim to fine-tune the GBM model's hyperparameters, potentially unlocking superior predictive performance compared to traditional, manually-tuned ML models.

This research is particularly timely given the growing emphasis on personalized medicine and the need for accurate risk stratification in HF management [13]. As healthcare systems worldwide grapple with resource allocation and treatment prioritization, especially in the wake of global health crises, refined prognostic tools could play a pivotal role in optimizing patient care pathways [14].

A. RESEARCH QUESTIONS

Our investigation seeks to address the following key questions:

1) How effective are optimized machine learning algorithms in predicting the survival of heart failure patients?
2) Which clinical and demographic factors most significantly impact the survival predictions for heart failure patients?
3) Does optimization improve the performance of machine learning models in heart failure survival prediction?

By exploring these questions, we aim to contribute to the evolving landscape of ML-assisted clinical decision support, potentially offering clinicians a more refined tool for prognostication in heart failure management.

B. KEY CONTRIBUTIONS

The key contributions of this work are summarized as follows:

1) **Introduction of AIW-PSO and GBM Combination:** This study introduces the novel combination of AIW-PSO and GBM for optimizing heart failure prediction models, demonstrating its potential to improve model performance by effectively tuning hyperparameters.
2) **Performance Across Balanced and Imbalanced Datasets:** The model's performance is explored across both balanced and imbalanced datasets, showcasing its practical utility in real-world applications, particularly in dealing with class imbalance issues that are common in medical datasets.
3) **Identification of Key Predictors:** Through feature selection techniques, the study identifies critical predictors, such as ejection fraction and serum creatinine, which provide valuable insights for clinical decision-making and contribute to the model's high accuracy.

Heart failure (HF) is a major public health concern that affects millions of people worldwide, resulting in high morbidity and mortality. The accurate prediction of survival in patients with heart failure is critical for guiding clinical practice. This study demonstrates a novel application of adaptive inertia weight particle swarm optimization (AIW-PSO) in conjunction with a Gradient Boosting Machine (GBM) for model performance improvement in prediction tasks. The proposed methodology shows the potential for improving the accuracy of models and providing meaningful clinical applications.

II. LITERATURE REVIEW

Heart failure (HF) is a critical public health issue, affecting approximately 26 million people worldwide and contributing significantly to morbidity and mortality rates [2]. The complex nature of HF, characterized by various etiologies and comorbidities, necessitates advanced predictive modeling to enhance clinical outcomes and guide treatment strategies [15]. The evolution of predictive modeling in healthcare, particularly in the realm of heart failure, has been marked by a shift from traditional statistical methods to more sophisticated machine learning approaches [16]. Accurate survival prediction in HF patients can enable clinicians to stratify risk effectively, optimize treatment plans, and allocate resources more efficiently, potentially improving patient outcomes and quality of life [17].

Additionally, Reinforcement learning (RL) has emerged as a powerful approach for optimizing decision-making in healthcare. Recent studies have demonstrated the potential of RL in learning treatment policies for critical conditions such as sepsis. However, deploying RL in healthcare settings requires robust model selection frameworks to address challenges like overfitting and computational complexity, particularly in offline settings. A recent work [36] investigates a practical model selection pipeline for offline RL using off-policy evaluation (OPE) methods. The study highlights Fitted Q Evaluation (FQE) as the most effective method for validation ranking, albeit at high computational costs, and proposes a two-stage approach to balance ranking accuracy and efficiency.

The relevance of such RL-based approaches aligns with this journal's focus on advancing machine learning applications for impactful real-world problems. While our work focuses on supervised learning using Gradient Boosting Machines (GBM) and AIW-PSO to classify heart failure patients, incorporating RL approaches like OPE could enable the development of adaptive and dynamic treatment policies in future studies. These advancements would not only expand the scope of predictive modeling but also enhance the practical utility of ML frameworks in clinical settings.

The application of machine learning (ML) algorithms in healthcare, especially in predicting patient survival and disease outcomes, has gained substantial momentum in recent years. Various studies have demonstrated the potential of these techniques to outperform traditional statistical methods in predicting heart failure outcomes. Awan et al. [18] utilized a combination of ML algorithms, including Support Vector Machines (SVM) and Random Forests (RF), to predict mortality in heart failure patients. Their study demonstrated superior predictive performance compared to conventional risk scores, highlighting the potential of ML in capturing complex interactions between clinical variables. In another significant study, Panahiazar et al. [19] developed a deep learning model for predicting 30-day readmission in heart failure patients. Their approach, which incorporated temporal trends in lab results and vital signs, achieved higher accuracy than traditional logistic regression models, underscoring the ability of advanced ML techniques to leverage complex, time-dependent data. Mortazavi et al. [20] compared various machine learning algorithms for predicting 1-year mortality in heart failure patients, finding that ensemble methods like Random Forests and Gradient Boosting Machines outperformed both traditional regression models and individual ML algorithms. This study emphasized the importance of algorithm selection and feature engineering in developing effective predictive models.

As the complexity of machine learning models increases, the need for effective optimization strategies [37], [38] becomes more pronounced. Hyperparameter optimization plays a crucial role in enhancing the performance of these models, and various techniques have been explored in the context of heart failure prediction. Bagheri et al. [21] employed a Genetic Algorithm (GA) for feature selection and hyperparameter tuning in their heart failure prediction model. Their approach, which combined GA with a Support Vector Machine classifier, demonstrated how optimization techniques can significantly improve model performance by identifying the most relevant predictors and optimal model configurations. In a novel approach, Beunza et al. [22] utilized Bayesian optimization for hyperparameter tuning in their ensemble model for predicting in-hospital mortality in heart failure patients. Their study showcased how advanced optimization techniques can enhance model performance while reducing computational overhead compared to traditional grid search methods. Alaa et al. [23] introduced an automated machine learning framework for clinical prediction tasks, including heart failure outcomes. Their approach, which used multi-armed bandits for model selection and Bayesian optimization for hyperparameter tuning, demonstrated state-of-the-art performance across various clinical prediction tasks.

Class imbalance presents a significant challenge in predictive modeling for heart failure, as mortality and adverse events are often relatively rare occurrences in clinical datasets. This imbalance can lead to biased models that perform poorly on the minority class, which is typically the class of greatest clinical interest. Choi et al. [24] addressed class imbalance in their study of heart failure readmission prediction by employing the Synthetic Minority Over-sampling Technique (SMOTE). Their approach, which combined SMOTE with ensemble learning methods, demonstrated improved predictive performance for the minority class without sacrificing overall model accuracy. Zahid et al. [25] explored various resampling techniques, including SMOTE and Adaptive Synthetic (ADASYN) sampling, in conjunction with different machine learning algorithms for heart failure prediction. Their comprehensive comparison provided insights into the effectiveness of different approaches to handling class imbalance in clinical prediction tasks. In a different approach, Guidi et al. [26] utilized cost-sensitive learning to address class imbalance in their study on predicting heart failure decompensation. By assigning higher misclassification costs to the minority class, they were able to improve the model's performance on this clinically important group without explicit resampling of the dataset.

III. METHOD

This study focuses on developing a robust machine learning model to predict heart failure survival, integrating various advanced techniques, including data preprocessing, feature selection, handling imbalanced data, and hyperparameter optimization using a nature inspired improved optimization algorithm called Particle Swarm Optimization Algorithm. Metaheuristic optimization algorithms, such as AIW-PSO (Adaptive Inertia Weight Particle Swarm Optimization), have proven to be highly effective [29] in improving the performance of machine learning models. These optimizers work by fine-tuning hyperparameters, which are often difficult to manually adjust, to enhance model accuracy, reduce overfitting, and improve overall model generalization. AIW-PSO and similar algorithms explore the search space intelligently to find optimal solutions without getting trapped in local optima. These steps are crucial for improving the accuracy and reliability of predictions. Figure 1 explains the overall architecture of the study.

A. DATA COLLECTION AND TRANSFORMATION

The dataset used in this study is the *Heart Failure Clinical Records* obtained from the Kaggle platform, containing 299 records with 13 clinical features. The primary objective is to predict the target variable *DEATH_EVENT*, which signifies whether a patient survived heart failure during the follow-up period. Features such as age, ejection fraction, serum creatinine, and high blood pressure are considered crucial indicators for survival prediction.

Data preprocessing included handling missing values, standardizing features using Standard Scaler, and splitting the dataset into training and testing sets with an 80:20 ratio. Standardization was applied to bring all features to a uniform scale, ensuring that algorithms sensitive to feature scaling, such as Support Vector Machines (SVM), perform optimally. Given the imbalanced nature of the dataset, with the majority of patients surviving, the Synthetic Minority Over-sampling Technique (SMOTE) was implemented. SMOTE works by generating synthetic samples for the minority class, in this case, death events, through interpolation between nearest neighbors in the feature space. This process ensures a balanced training dataset, improving the model's ability to generalize to unseen data while preventing bias toward the majority class.

The goal of our research was to ascertain how long heart failure patients may survive. We have used a number of well-known machine learning techniques that have been enhanced by *AIW-PSO* (Adaptive Inertia Weight Particle Swarm Optimization) to achieve this. The outcomes of our computations utilizing the 13 and 7 best attributes are displayed. Furthermore, our findings suggest that the *SMOTE* method is the most beneficial inequality provision. Table 1 shows a summarized overview of the dataset.

B. FEATURE SELECTION

Feature selection was performed using SelectKBest and Chi-square tests to reduce dimensionality and retain only the most relevant characteristics for the prediction of heart failure survival. SelectKBest selects the best k features based on the ANOVA F value, while the Chi-square test evaluates the relationship between each feature and the target variable. The Chi-Square Test for Independence is used to determine whether there is a significant association between two categorical variables in a dataset. The Chi-square statistic is computed using the following formula:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}, \tag{1}$$

where O_i and E_i are the observed and expected frequencies, respectively.

The results revealed that features such as *age*, *serum creatinine*, and *ejection fraction* were among the most predictive of patient survival. Reducing the feature space through this process not only improves the computational efficiency of the models, but also enhances interpretability by focusing on the most significant predictors.

C. CORRELATION ANALYSIS OF FEATURES

To ensure an accurate correlation analysis between variables, we distinguished between continuous-continuous variable pairs and categorical-continuous variable pairs.

1) For **continuous-continuous pairs**, we employed the *Pearson correlation coefficient*, which measures the linear association between two continuous variables and is computed as:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}, \tag{2}$$

where x_i and y_i are the values of the two continuous variables, and $\bar{x}$ and $\bar{y}$ are their respective means.

2) For **categorical-continuous pairs** (e.g., binary variables like *anaemia*, *diabetes*, *sex*, and *smoking*), we applied the *Point-Biserial correlation coefficient*, which is appropriate for associations of categorical and continuous binary variables. The Point-Biserial correlation is computed as:

$$r_{pb} = \frac{\bar{X}_1 - \bar{X}_0}{s}\sqrt{\frac{n_1 n_0}{n^2}}, \tag{3}$$

where:

- $\bar{X}_1$ and $\bar{X}_0$ are the means of the continuous variable for the two binary classes (1 and 0),
- s is the standard deviation of the continuous variable,

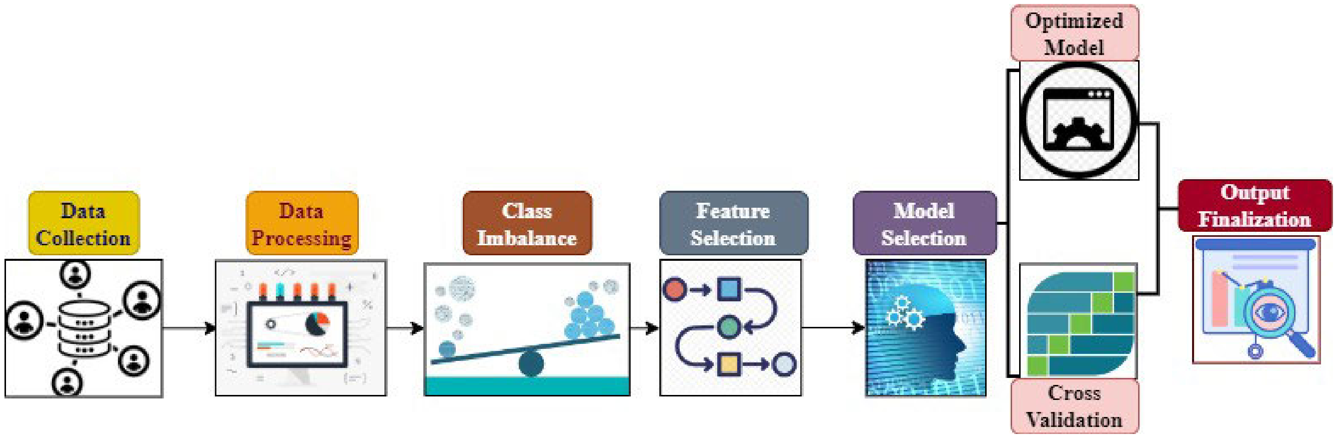

FIGURE 1. Procedural architecture of our study.

TABLE 1. Feature description for heart failure dataset.

Feature Name	Description	Type	Category	Range/Values
age	Age of the patient	Continuous (float)	Input	40-95 years
anaemia	Decrease of red blood cells or hemoglobin (1=yes, 0=no)	Categorical (int)	Input	0, 1
creatinine_phosphokinase	Level of the CPK enzyme in the blood (mcg/L)	Continuous (float)	Input	23 - 7861 mcg/L
diabetes	If the patient has diabetes (1=yes, 0=no)	Categorical (int)	Input	0, 1
ejection_fraction	Percentage of blood leaving the heart at each contraction	Continuous (float)	Input	14% - 80%
high_blood_pressure	If the patient has hypertension (1=yes, 0=no)	Categorical (int)	Input	0, 1
platelets	Platelets in the blood (kiloplatelets/mL)	Continuous (float)	Input	25.1 - 850.0 kiloplatelets/mL
serum_creatinine	Level of serum creatinine in the blood (mg/dL)	Continuous (float)	Input	0.5 - 9.4 mg/dL
serum_sodium	Level of serum sodium in the blood (mEq/L)	Continuous (float)	Input	113 - 148 mEq/L
sex	Gender of the patient (1=male, 0=female)	Categorical (int)	Input	0, 1
smoking	If the patient smokes (1=yes, 0=no)	Categorical (int)	Input	0, 1
time	Follow-up period (days)	Continuous (int)	Input	4 - 285 days
DEATH_EVENT	If the patient died during the follow-up period (1=yes, 0=no)	Categorical (int)	Output	0, 1

- n_1 and n_0 are the number of observations in each class,
- n is the total number of observations.

The correlation analysis results were visualized in a heatmap (Figure 2), where Pearson correlation was used for continuous pairs and Point-Biserial correlation for categorical-continuous pairs. This method ensures that the relationships between features are accurately measured based on their respective types, avoiding incorrect statistical assumptions.

The heatmap highlights the key relationships between the characteristics and the target variable (*DEATH_EVENT*). For example, *age*, *serum creatinine*, and *ejection fraction* show strong associations with the target variable. This approach not only enhances the rigor of the analysis, but also ensures methodological correctness.

D. CLASS ENCODING

Since some features in the data set were categorical, such as sex, the One-Hot Encoding [29] was applied to convert these categorical variables into numerical binary columns. For instance, the sex variable was split into two binary columns: `sex_male` and `sex_female`. This encoding ensures that machine learning models can correctly interpret categorical features without introducing ordinal biases. One-Hot Encoding is particularly effective when dealing with non-ordinal features and ensures compatibility with algorithms such as Random Forest and Support Vector Machines (SVM).

E. HANDLING IMBALANCED DATA

The dataset presented a significant class imbalance, with the majority of patients surviving and a smaller proportion experiencing death events. To address this, SMOTE (Synthetic Minority Over-sampling Technique) [30], [31], [32] was employed. SMOTE generates synthetic samples for the minority class by selecting examples from the minority class and interpolating between their nearest neighbors in the feature space. This technique effectively balances the dataset, enabling the model to learn more effectively from the minority class. After applying SMOTE, the class distribution was more balanced, leading to improved generalization and prediction performance on the minority class (death events)

In each category, there are two different kinds of information altogether. Asymmetric representation of the target two classes with attributes 13 and 7 characterizes one kind, whereas a balanced distribution of the two classes characterizes the other, Figure: 3 explains.

F. HYPERPARAMETER OPTIMIZATION

To improve the performance of the machine learning models, hyperparameter optimization was conducted using the Mealpy [27] framework. Instead of using traditional

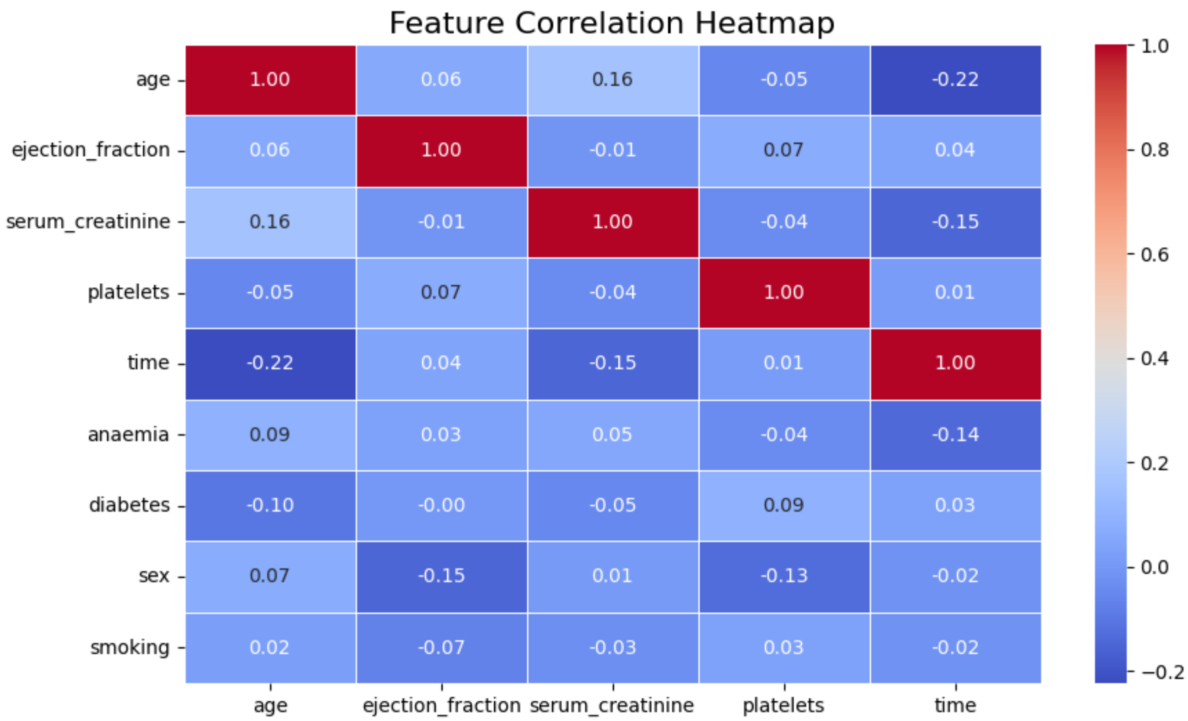

FIGURE 2. Feature Correlation Heatmap showing Pearson correlation for continuous-continuous pairs and Point-Biserial correlation for binary categorical-continuous pairs.

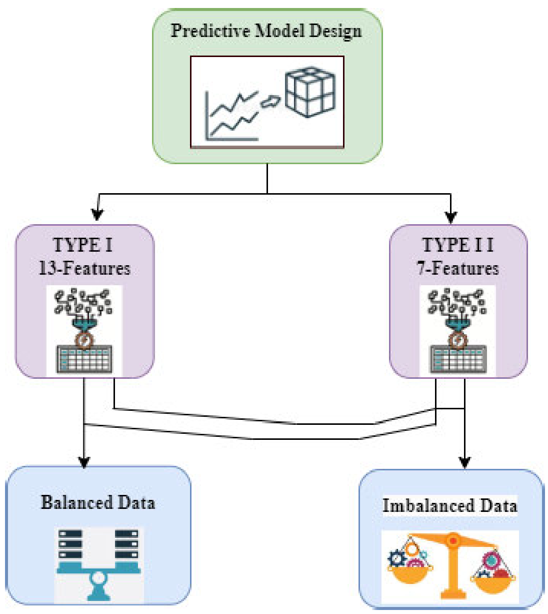

FIGURE 3. Proposed types of predictive model designs.

optimization methods, this study employed Adaptive Inertia Weight Particle Swarm Optimization (AIW-PSO) [28], a variant of the standard Particle Swarm Optimization (PSO) algorithm. AIW-PSO is known for its enhanced exploration-exploitation balance through adaptive adjustments of the inertia weight, which improves convergence speed and avoids local optima more effectively than the original PSO. The Mealpy library was chosen due to its flexibility and robust implementation of various metaheuristic algorithms, including AIW-PSO. This optimization was applied to several machine learning models to fine-tune their hyperparameters, Table 2 describes the hyperparameters optimized by AIW-PSO.

G. AIW-PSO ALGORITHM

The Adaptive Inertia Weight Particle Swarm Optimization (AIW-PSO) algorithm enhances the classical PSO by adaptively adjusting the inertia weight w over iterations, effectively balancing exploration and exploitation. The update equation for each particle's velocity and position is given by:

Velocity update:

$$v_i^{t+1} = w \cdot v_i^t + c_1 \cdot r_1 \cdot (p_{\text{best},i} - x_i^t) + c_2 \cdot r_2 \cdot (g_{\text{best}} - x_i^t) \tag{4}$$

where:

- v_i^{t+1} is the velocity of particle i at time step $t+1$,
- w is the inertia weight, which is adaptively adjusted using AIW strategy,
- c_1 and c_2 are acceleration coefficients,
- r_1 and r_2 are random values in the range [0,1],
- $p_{best,i}$ is the personal best position of particle i,
- g_{best} is the global best position,
- x_i^t is the position of particle i at time step t.

TABLE 2. Hyperparameters optimized by AIW-PSO for various algorithms.

Model	Hyperparameters Optimized by AIW-PSO	Explanation of Hyperparameters
Random Forest (RF)	n_estimators, max_depth, min_samples_split	- n_estimators: Controls the number of trees. - max_depth: Limits how deep each tree grows. - min_samples_split: Minimum samples required to split a node.
Support Vector Classifier (SVC)	C, gamma, kernel	- C: Balances classification accuracy and margin width. - gamma: Defines the influence of a single training point. - kernel: Specifies the type of kernel function.
AdaBoost	n_estimators, learning_rate	- n_estimators: Number of boosting rounds. - learning_rate: Controls the contribution of each weak learner. - algorithm: Type of boosting (SAMME, SAMME.R).
Gradient Boosting Machine (GBM)	n_estimators, learning_rate, max_depth	- n_estimators: Number of boosting stages. - learning_rate: Shrinks the contribution of each stage. - max_depth: Limits the depth of each tree.
Stochastic Gradient Descent (SGD)	alpha, eta0, penalty	- alpha: Controls regularization to prevent overfitting. - eta0: Sets the initial learning rate. - penalty: Specifies the regularization type (l1, l2, elasticnet).

Position update:

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{5}$$

In AIW-PSO, the inertia weight w is updated as:

$$w = w_{\max} - \left(\frac{w_{\max} - w_{\min}}{iter_{\max}}\right) \cdot iter \tag{6}$$

where $iter$ is the current iteration, $iter_{max}$ is the maximum number of iterations, w_{max} and w_{min} are predefined maximum and minimum inertia weights. This adaptive adjustment helps the particles to explore the search space more thoroughly at the beginning and gradually shift to exploitation in later stages of the search.

H. INTEGRATION WITH GRADIENT BOOSTING MACHINE (GBM)

In this study, AIW-PSO was integrated with Gradient Boosting Machine (GBM) to optimize its hyperparameters, improving the model's ability to predict heart failure survival. GBM is a powerful ensemble machine learning algorithm that builds sequential trees where each subsequent tree attempts to correct the errors of its predecessor. The hyperparameters optimized using AIW-PSO include the number of estimators $n_{estimators}$, learning rate η, and maximum tree depth max_depth, which significantly influence the model's performance and its ability to generalize to unseen data.

The AIW-PSO algorithm efficiently searches for the best combination of GBM hyperparameters by minimizing the error function over iterations. The objective function used in the optimization process was based on the cross-validated accuracy of the GBM model. Through this combined approach, the proposed model achieved enhanced prediction accuracy and model stability compared to traditional methods, as demonstrated by the experimental results. This highlights the benefit of using meta-heuristic optimization methods like AIW-PSO to improve the effectiveness of machine learning models in complex real-world applications like heart failure survival prediction.

1) OBJECTIVE FUNCTION

The objective function for hyperparameter optimization in GBM is defined as minimizing the classification error, which can be represented mathematically as:

$$f(\mathbf{x}) = 1 - \text{Accuracy}(n_{\text{estimators}}, \text{learning\_rate}, \text{max\_depth}) \tag{7}$$

where:

- $n_{\text{estimators}}$ is the number of boosting stages,
- learning_rate controls the contribution of each tree,
- max_depth limits how deep each tree can grow.

2) STEPS FOR AIW-PSO AND GBM INTEGRATION

The integration of AIW-PSO with GBM involves the following steps and Figure:4 shows the flow diagram.

1) **Define the search space:** Specify the hyperparameters of the GBM model to be optimized, represented as:

$$\mathbf{x} = [n_{\text{estimators}}, \text{learning\_rate}, \text{max\_depth}]$$

2) **Initialize particles:** Each particle represents a potential solution in the search space, containing a set of hyperparameters for the GBM model.
3) **Update particle velocities and positions:** Use the AIW-PSO equations (4) and (5) to update the velocities and positions of the particles iteratively.
4) **Evaluate GBM performance:** For each particle's position, evaluate the GBM model's performance using the defined objective function.
5) **Update personal and global best positions:** Keep track of the best-performing solutions encountered by each particle as well as the global best solution.
6) **Terminate the process:** The optimization process continues until the stopping criteria are met (e.g., reaching a maximum number of iterations).
7) **Train the final GBM model:** Using the optimized hyperparameters from the best solution, train the final GBM model and evaluate its performance on the test set.

The overall process has been decomposed into three main phases shown in Figure 5. The objective function for

Algorithm 1 Calculation of Information Criteria (IC) Values

Input: n: Number of data points, k: Number of model parameters, LL: Log-likelihood value.
Output: IC values (AIC, BIC, HQIC, AICc).
begin
  Calculate AIC: $AIC = 2k - 2LL$;
  Calculate BIC: $BIC = k\log(n) - 2LL$;
  Calculate HQIC: $HQIC = 2k\log(\log(n)) - 2LL$;
  Calculate AICc: $AICc = AIC + \frac{2k(k+1)}{n-k-1}$;
  return AIC, BIC, HQIC, AICc;

AIW-PSO was defined as minimizing classification error (1 - accuracy). The AIW-PSO algorithm explores potential hyperparameter solutions by simulating particles moving through a solution space, updating their positions based on both individual and global best solutions. By adaptively adjusting the inertia weight, AIW-PSO was able to strike an optimal balance between exploration and exploitation, leading to superior hyperparameter tuning results compared to the original PSO.The hyperparameter optimization process led to substantial performance improvements across all models, with the gradient boosting model achieving the highest accuracy (93.84%). AIW-PSO's adaptive behavior allowed it to converge more quickly and find better solutions, resulting in more efficient model training and higher predictive accuracy.

IV. MODEL EVALUATION AND RESULTS ANALYSIS

Following hyperparameter optimization, the models were evaluated using key performance metrics, including accuracy, precision, recall, F1-score, and the area under the ROC curve (AUC).

In machine learning classification tasks, by selecting appropriate evaluation metrics is critical to accurately assess model performance. While Information Criteria (AIC, SBIC, HQIC, AICc) are well-suited for statistical model selection, their utility is more aligned with parametric models such as regression or likelihood-based models. In contrast, this study focuses on evaluating and optimizing the predictive performance of machine learning classifiers, which require metrics that are tailored for classification outcomes.

A. RATIONALE FOR ACCURACY, PRECISION, RECALL, AND F1-SCORE

To evaluate the proposed Gradient Boosting Machine (GBM) model optimized using Adaptive Inertia-Weight Particle Swarm Optimization (AIW-PSO), we employed the following widely accepted metrics:

- **Accuracy:** Represents the proportion of correctly predicted instances among the total instances. Accuracy is a fundamental measure for assessing overall model performance.
- **Precision:** Evaluates the ratio of true positive predictions to the total positive predictions, reflecting the ability of the model to avoid false positives.
- **Recall (Sensitivity):** Measures the ratio of true positives to all actual positives, indicating the model's ability to identify all positive cases.
- **F1-Score:** The harmonic mean of Precision and Recall, providing a balanced metric when both false positives and false negatives are of concern.

These metrics are particularly relevant for imbalanced datasets, such as the heart failure survival dataset used in this study. Accuracy alone may fail to provide an accurate assessment in imbalanced scenarios; therefore, we incorporated Precision, Recall, and F1-Score to provide a more comprehensive evaluation of the model's performance.

B. INTEGRATION OF INFORMATION CRITERIA (IC) VALUES FOR CLASSIFICATION

To enhance the model's performance and ensure minimized Information Criterion (IC) values, several strategies can be employed. First, feature selection and regularization techniques can be utilized to reduce dataset dimensionality and limit model complexity, which directly impacts IC calculations by lowering the number of parameters (k). Regularization methods such as L1 (Lasso) or L2 (Ridge) can be introduced during training to mitigate overfitting, further contributing to robustness.

Second, hyperparameter tuning should be performed with greater granularity, particularly refining ranges and step sizes for parameters like *n_estimators*, *learning_rate*, and others. This includes tuning additional regularization parameters like *min_samples_leaf*, *subsample*, and *max_features*. Incorporating log-loss as an optimization target can also prove beneficial, as it is inherently linked to IC computations; minimizing log-loss directly reduces IC values.

Advanced optimizers, such as hybrid approaches combining AIW-PSO with local search methods, can further refine parameter optimization for precision. Lastly, dataset balancing techniques, such as SMOTE, should be revisited to avoid introducing synthetic noise that could inflate IC values, ensuring a cleaner dataset for robust model evaluation.

The algorithm of IC values calculation used in this study shown in Algorithm 1 and the explanation showed in Table 3. By including IC values in the analysis, researchers can assess model robustness more comprehensively, as these metrics combine model fit and complexity, offering insights into the trade-offs between predictive performance and overfitting.

C. PERFORMANCE EVALUATION

In this study, the performance of various machine learning algorithms for heart failure survival prediction was evaluated, categorizing the results into two types based on the number of features used. Type I refers to the analysis using 13 features under both imbalanced and balanced conditions, while Type II refers to the results obtained using 7 features under the same conditions. The models assessed include Random

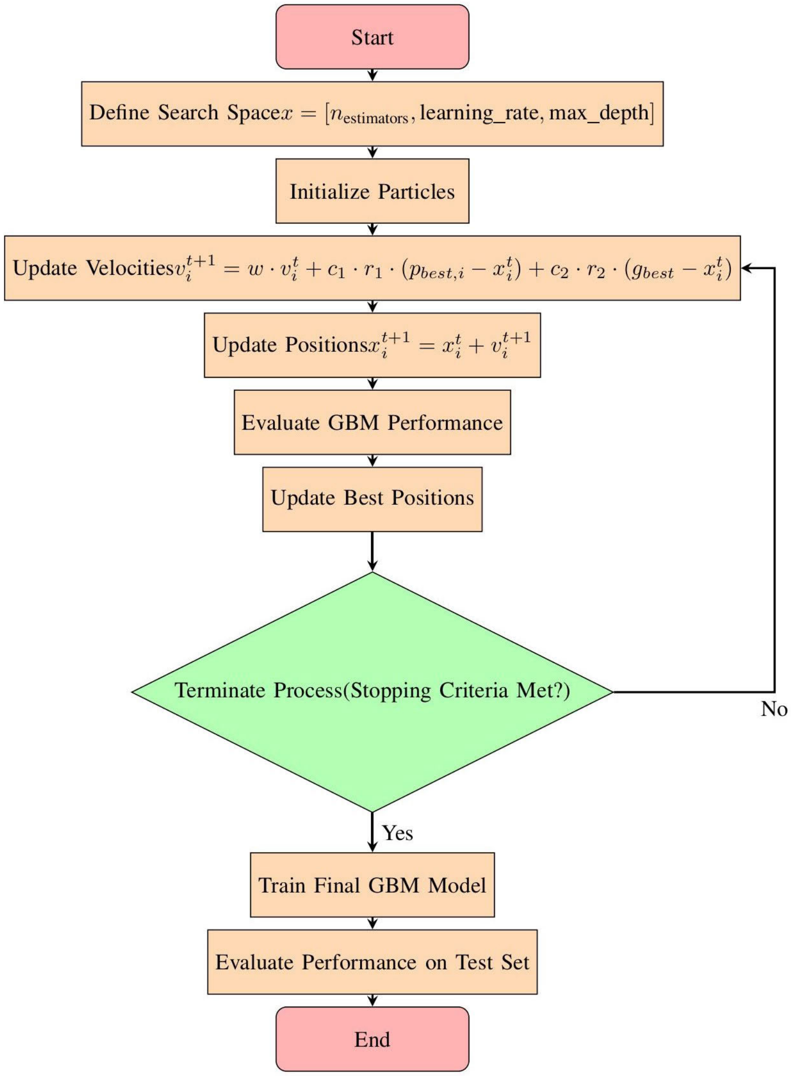

FIGURE 4. GBM based AIW-PSO Work-flow diagram.

TABLE 3. Explanation of algorithm for calculating information criteria (IC) values.

Aspect	Description
Inputs	• n: Number of data points in the test set. • k: Number of model parameters, including regularization terms and interactions. • LL: Log-likelihood derived from log-loss or other likelihood-based metrics.
Steps	1. Calculate Akaike Information Criterion (AIC): Balances model complexity (k) and fit (LL). 2. Calculate Bayesian Information Criterion (BIC): Introduces a stronger penalty for complexity relative to the dataset size (n). 3. Calculate Hannan-Quinn Information Criterion (HQIC): Applies a logarithmic penalty on the sample size. 4. Calculate Corrected AIC (AICc): Adjusts AIC to address small sample sizes, mitigating overestimation of model fit.
Outputs	Returns minimized IC values (AIC, BIC, HQIC, AICc) to identify the optimal model balancing predictive accuracy and parsimony.

Forest (RF), Support Vector Classifier (SVC), AdaBoost, Gradient Boosting Machine (GBM), and Stochastic Gradient Descent (SGD).

1) TYPE I: 13 FEATURES - BALANCED AND IMBALANCED

For Type I, the results in Figure 6 indicated that using all 13 features with SMOTE significantly improved the model performance metrics. The Gradient Boosting Machine (GBM) achieved an accuracy of 87.79% and an F1 score of 0.89. The Random Forest (RF) model performed competitively with an accuracy of 84.34% and an F1 score of 0.85. In contrast, SVC exhibited lower performance, attaining an accuracy of 81.12%. Notably, when comparing imbalanced and balanced data, the implementation of SMOTE markedly enhanced the models' performance, showcasing its effectiveness in addressing the class imbalance issue prevalent in the dataset.

Table 4 presents a comprehensive performance comparison of optimized machine learning algorithms, incorporating AIW-PSO, on both imbalanced and balanced datasets using

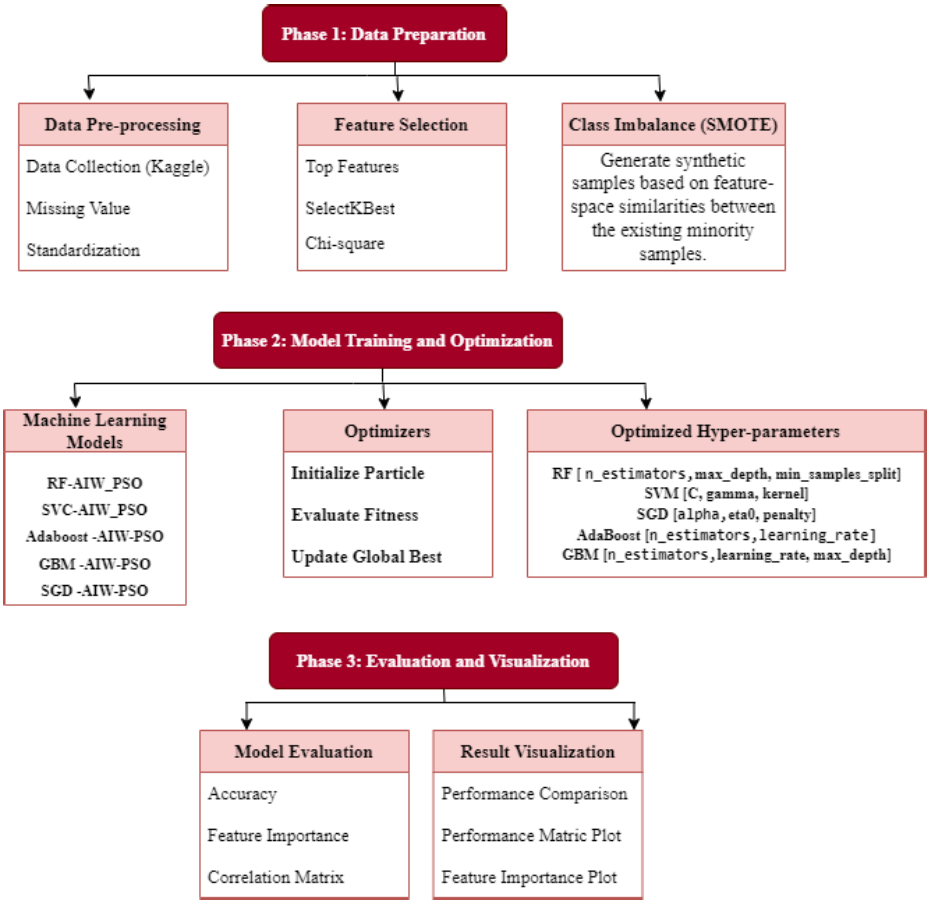

FIGURE 5. Three phases of methodology.

various metrics, including Accuracy, F1 Score, Precision, Recall, and Information Criteria (IC) values (AIC, BIC, HQIC, and AICc). Across the imbalanced dataset, the GBM-AIW_PSO algorithm achieved the highest accuracy (83.58%) and F1 Score (0.85), coupled with the lowest AIC (9.33), BIC (21.83), and HQIC (13.73), demonstrating superior robustness and predictive performance. Conversely, the SGD-AIW_PSO algorithm showed significantly lower accuracy (42.67%) and higher IC values, indicating poorer performance. On the balanced dataset, GBM-AIW_PSO continued to excel with the highest accuracy (87.79%), F1 Score (0.89), and the lowest AIC (9.23), highlighting its robustness in handling balanced datasets. Other algorithms, such as Adaboost-AIW_PSO and RF-AIW_PSO, also performed competitively but were outperformed by GBM-AIW_PSO in most metrics. The balanced dataset notably improved overall performance metrics across all algorithms, emphasizing the importance of dataset balancing in predictive modeling. These results underscore the robustness and reliability of GBM-AIW_PSO as a state-of-the-art approach for both imbalanced and balanced datasets.

2) TYPE II: 7 FEATURES - BALANCED AND IMBALANCED

For Type II, shown in the following Figure 7 and in tabular format in Table 5 utilizing a reduced set of 7 features yielded even more promising results. The GBM model achieved the highest accuracy of 93.84% with a robust F1 score of 0.95, indicating its superior predictive capability when fewer features were employed. Similarly, the Random Forest (RF) model demonstrated a significant improvement, reaching an accuracy of 88.11%. The results from the reduced feature set further highlight the importance of feature selection in improving model performance, as evidenced by the substantial gains in accuracy across all algorithms tested.

The implementation of AIW-PSO for hyperparameter optimization played a crucial role in enhancing the performance of the machine learning models in this study. By optimizing parameters such as n_estimators and learning_rate for GBM, and C and gamma for SVC, the models could achieve higher accuracy levels and better generalization. For instance, using AIW-PSO optimization led to the GBM model's impressive accuracy of 93.84% in Figure 9, illustrating how effective hyperparameter tuning is in optimizing machine learning algorithms. The results underscore the significance of employing advanced optimization techniques like AIW-PSO to boost predictive performance and ensure more reliable survival analysis in clinical applications. ALL the comparison shown in bar-chart in the Figure 8.

V. DISCUSSION

This study effectively addresses the critical research questions regarding the survival prediction of heart failure

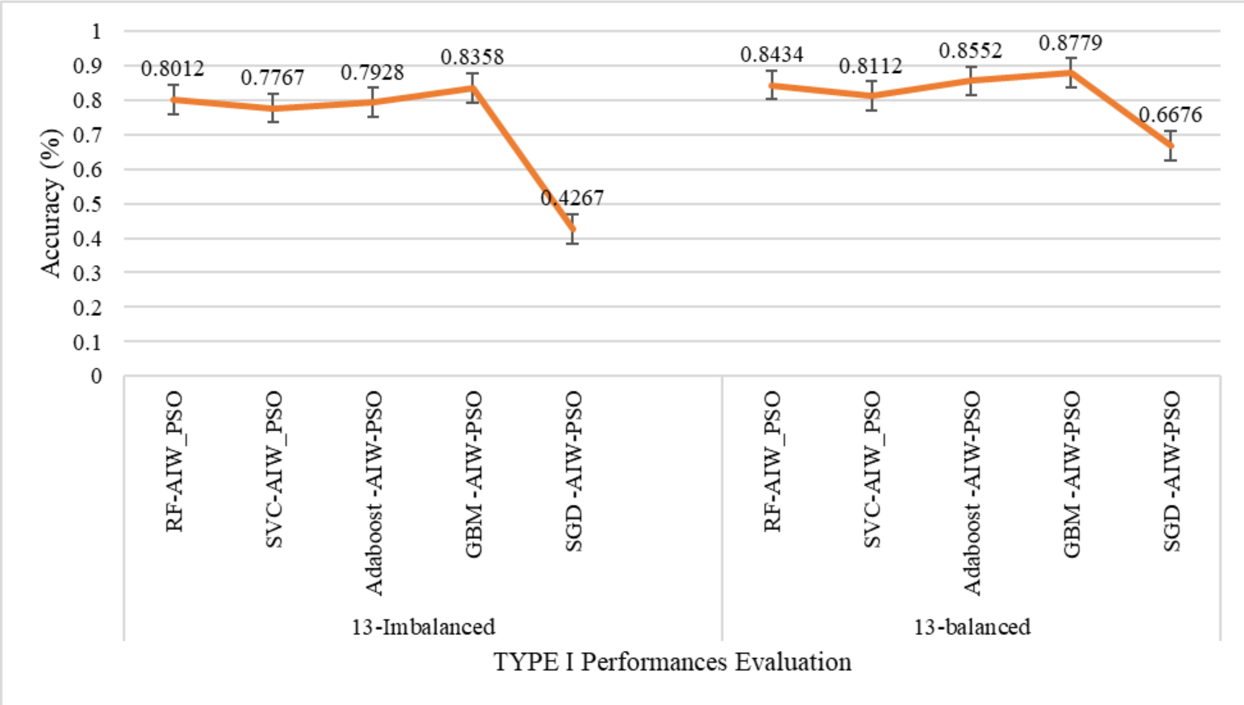

FIGURE 6. Type I performance evaluation.

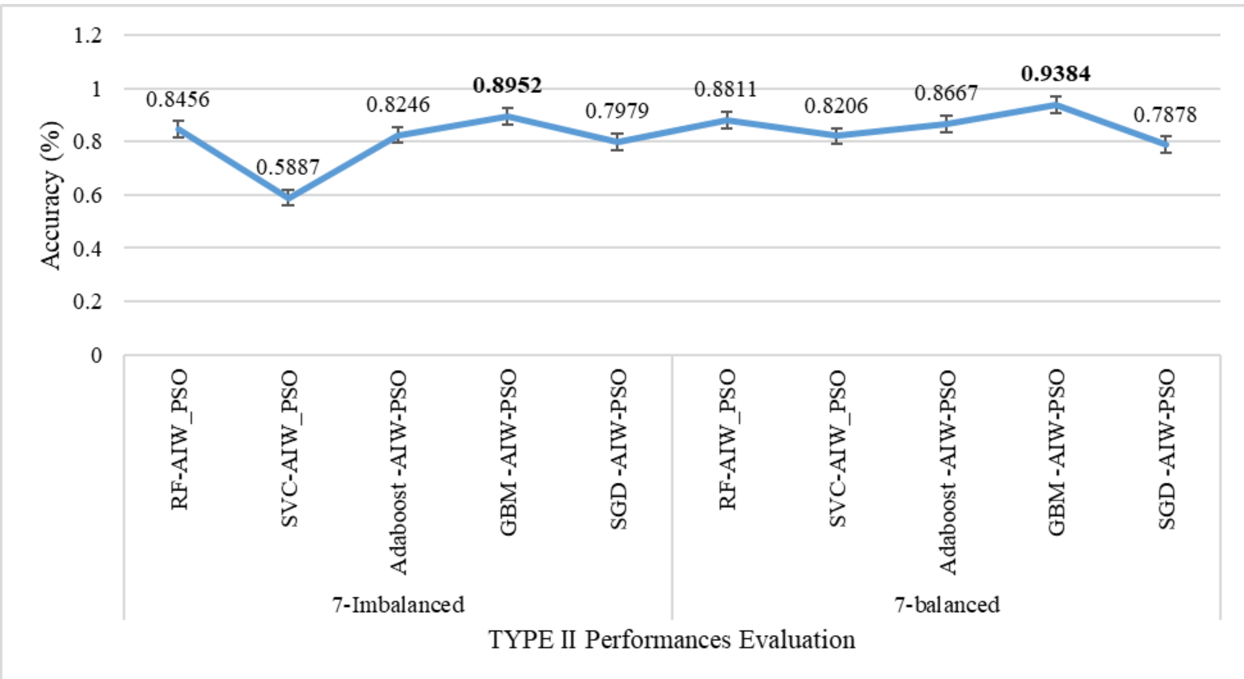

FIGURE 7. Type II performance evaluation.

patients through the application of optimized machine learning algorithms. The results demonstrate that optimized models, particularly Gradient Boosting Machine (GBM) and Random Forest (RF), significantly enhance predictive accuracy, achieving up to 93.84% (by GBM) which is superior to [33] accuracy when utilizing advanced optimization techniques such as AIW-PSO. The analysis identifies key clinical and demographic factors that influence survival

TABLE 4. Type I performance comparison of optimized ml algorithms on imbalanced and balanced datasets with IC values.

Dataset	Algorithm	Accuracy (%)	F1 Score	Precision	Recall	AIC	BIC	HQIC	AICc
13-Imbalanced	RF-AIW_PSO	80.12	0.81	0.82	0.81	10.42	22.92	17.83	10.63
	SVC-AIW_PSO	77.67	0.78	0.78	0.77	10.50	21.00	17.90	10.70
	Adaboost-AIW_PSO	79.28	0.79	0.80	0.78	10.47	21.97	17.88	10.68
	GBM-AIW_PSO	83.58	0.85	0.83	0.84	9.33	21.83	13.73	10.53
	SGD-AIW_PSO	42.67	0.46	0.43	0.42	11.55	30.06	18.96	11.76
13-Balanced	RF-AIW_PSO	84.34	0.85	0.84	0.84	10.33	2.83	17.73	10.53
	SVC-AIW_PSO	81.12	0.83	0.82	0.82	10.37	28.87	17.78	10.58
	Adaboost-AIW_PSO	85.52	0.84	0.86	0.86	10.35	28.85	17.75	10.55
	GBM-AIW_PSO	87.79	0.89	0.88	0.89	9.23	20.74	13.64	9.44
	SGD-AIW_PSO	66.76	0.65	0.68	0.68	10.86	29.36	18.27	11.07

TABLE 5. Type II performance comparison of optimized ML algorithms on imbalanced and balanced datasets.

Dataset	Algorithm	Accuracy (%)	F1 Score	Precision	Recall
7-Imbalanced	RF-AIW_PSO	84.56	0.86	0.85	0.85
	SVC-AIW_PSO	58.87	0.59	0.58	0.57
	Adaboost-AIW_PSO	82.46	0.83	0.82	0.82
	GBM-AIW_PSO	89.52	0.89	0.90	0.89
	SGD-AIW_PSO	79.79	0.77	0.79	0.78
7-Balanced	RF-AIW_PSO	88.11	0.88	0.87	0.88
	SVC-AIW_PSO	82.06	0.83	0.82	0.82
	Adaboost-AIW_PSO	86.67	0.87	0.87	0.86
	GBM-AIW_PSO	93.84	0.95	0.94	0.94
	SGD-AIW_PSO	78.78	0.79	0.78	0.79

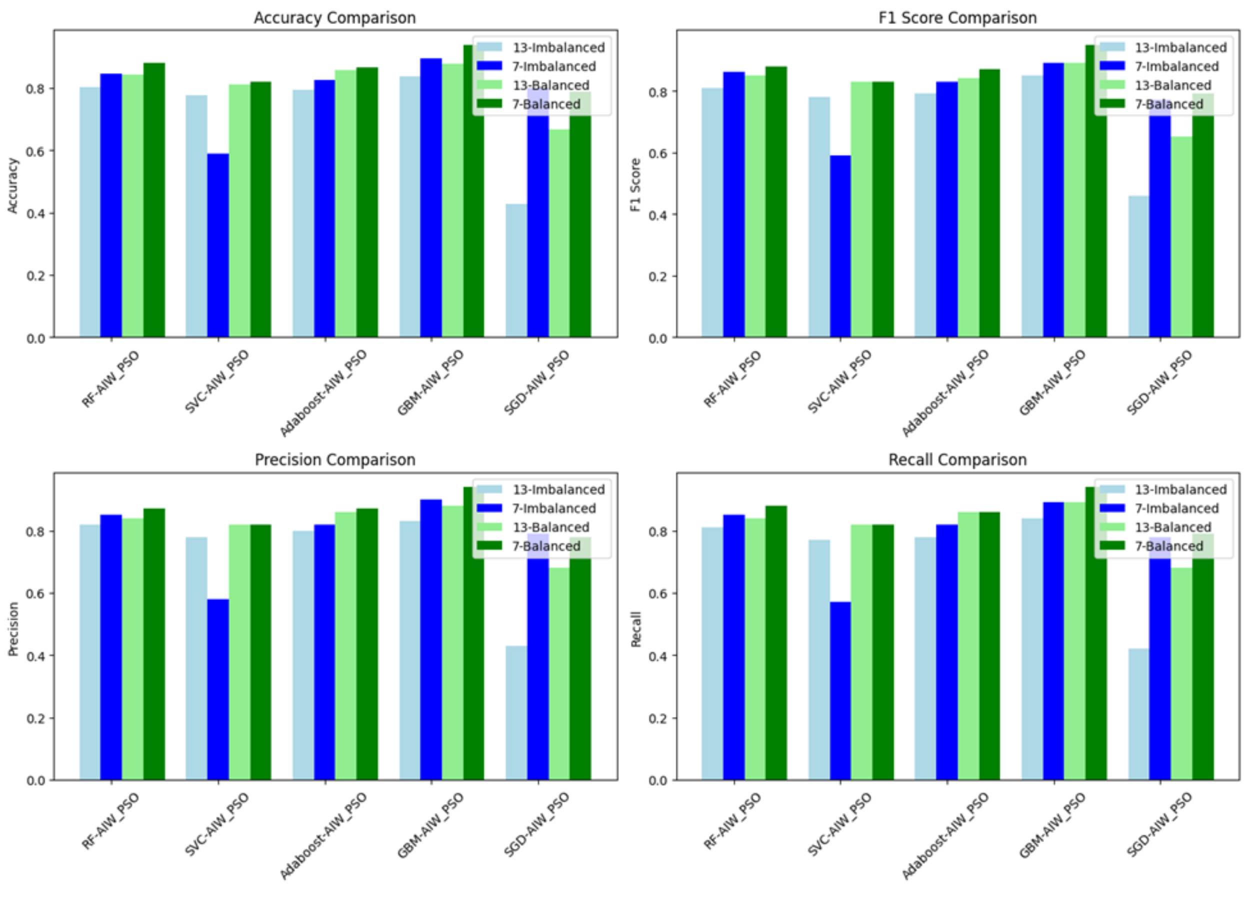

FIGURE 8. Accuracy comparison among all algorithms.

predictions, including age, serum creatinine levels, and ejection fraction, thereby providing valuable insights into the attributes that clinicians should monitor closely. Furthermore, the optimization process, which fine-tunes hyperparameters of various algorithms like SVC, AdaBoost, and Stochastic Gradient Descent (SGD), has been shown to improve model performance consistently. The comprehensive analyses, including the handling of class imbalance using SMOTE

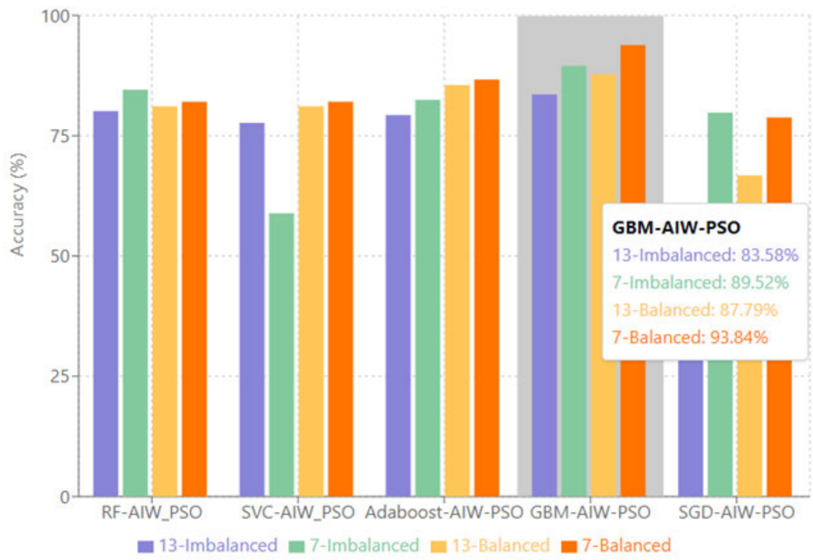

FIGURE 9. Highlighted GBM-AIW-PSO performances.

and feature selection methods, underscore the importance of employing sophisticated machine learning techniques to achieve reliable survival predictions, ultimately contributing to better clinical decision-making and patient outcomes in heart failure management.

Recent advancements in machine learning, such as the development of FLUID-GPT, have demonstrated the potential of transformer-based architectures for predictive modeling in complex systems. FLUID-GPT, a hybrid model combining Generative Pre-Trained Transformer 2 (GPT-2) with a Convolutional Neural Network (CNN), has been applied to predict particle trajectories and surface erosion patterns in industrial-scale systems [35]. By leveraging information from initial conditions such as particle size, inlet speed, and pressure, FLUID-GPT achieves a 54% reduction in mean squared error and 70% faster training times compared to traditional BiLSTM approaches. These advancements illustrate the growing role of generative transformer-based models in replacing computationally expensive simulations, particularly for dynamic and time-series predictions. While this study focuses on GBM with AIW-PSO for the classification of heart failure patients, the integration of transformer-based architectures like FLUID-GPT could be a valuable direction for future exploration, especially for clinical datasets requiring sophisticated time-series modeling.Future work may consider adapting such architectures to clinical datasets, enabling the incorporation of contextual and sequential data for more accurate predictions in healthcare applications.

VI. CONCLUSION AND FUTURE WORK

This study evaluates the performance of various machine learning models for the prediction of heart failure survival, incorporating feature selection and class balancing using SMOTE. Among the models tested, Random Forest (RF), Support Vector Classifier (SVC), AdaBoost, Gradient Boosting Machine (GBM), and Stochastic Gradient Descent (SGD), GBM achieved the highest accuracy of 93.84% with SMOTE, demonstrating its robustness in handling imbalanced datasets. Feature selection significantly impacted model performance, improving some models while reducing the effectiveness of others. The study highlights the critical role of hyperparameter tuning through AIW-PSO, which enhanced model performance.

Additionally, information criteria such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were incorporated to refine model selection, ensuring an optimal balance between predictive accuracy and complexity. GBM exhibited the lowest AIC and BIC values, reinforcing its superiority. The findings emphasize that integrating advanced optimization techniques and addressing class imbalance can significantly enhance prediction outcomes. The inclusion of information criteria provides a rigorous model evaluation framework, contributing valuable insights into heart failure survival analysis and improving clinical decision-making.

Future research could focus on three key areas to improve predictive modeling for heart failure survival analysis. Firstly, expanding the data set to include additional demographic and clinical variables, along with longitudinal data, could provide a more comprehensive understanding of the factors that influence survival, potentially improving the accuracy and interpretability of the model. Secondly, exploring advanced ensemble methods or hybrid models that combine the strengths of various algorithms, such as stacking or blending multiple machine learning techniques, could yield even better predictive performance. Lastly, investigating alternative optimization algorithms beyond AIW-PSO, such as Genetic Algorithms or Differential Evolution, may uncover novel hyperparameter configurations that enhance model robustness and adaptability to different datasets, ensuring that the predictions remain accurate in diverse clinical settings.

REFERENCES

[1] T. A. McDonagh et al., "2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure," *Eur. Heart J.*, vol. 42, no. 36, pp. 3599–3726, Sep. 2021, doi: 10.1093/eurheartj/ehab368.

[2] G. Savarese and L. H. Lund, "Global public health burden of heart failure," *Cardiac Failure Rev.*, vol. 3, no. 1, pp. 7–11, 2017, doi: 10.15420/cfr.2016:25:2.

[3] S. S. Virani et al., "Heart disease and stroke statistics—2021 update: A report from the American heart association," *Circulation*, vol. 143, no. 8, pp. e254–e743, Feb. 2021, doi: 10.1161/cir.0000000000000950.

[4] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New England J. Med.*, vol. 380, no. 14, pp. 1347–1358, Apr. 2019, doi: 10.1056/nejmra1814259.

[5] S. Angraal, B. J. Mortazavi, A. Gupta, R. Khera, T. Ahmad, N. R. Desai, D. L. Jacoby, F. A. Masoudi, J. A. Spertus, and H. M. Krumholz, "Machine learning prediction of mortality and hospitalization in heart failure with preserved ejection fraction," *JACC, Heart Failure*, vol. 8, no. 1, pp. 12–21, Jan. 2020, doi: 10.1016/j.jchf.2019.06.013.

[6] M. Feurer and F. Hutter, "Hyperparameter optimization," in *Automated Machine Learning*, F. Hutter, L. Kotthoff, and J. Vanschoren, Eds., Cham, Switzerland: Springer, 2019, pp. 3–33, doi: 10.1007/978-3-030-05318-5_1.

[7] D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: An overview," *Soft Comput.*, vol. 22, no. 2, pp. 387–408, Jan. 2018, doi: 10.1007/s00500-016-2474-6.

[8] Y. Xue, J. Jiang, B. Zhao, and T. Ma, "A self-adaptive artificial bee colony algorithm based on global best for global optimization," *Soft Comput.*, vol. 22, no. 9, pp. 2935–2952, May 2018, doi: 10.1007/s00500-017-2547-1.

[9] T. Ahmad, A. Munir, S. H. Bhatti, M. Aftab, and M. A. Raza, "Survival analysis of heart failure patients: A case study," *PLoS ONE*, vol. 12, no. 7, Jul. 2017, Art. no. e0181001, doi: 10.1371/journal.pone.0181001.

[10] A. Fernandez, S. Garcia, F. Herrera, and N. V. Chawla, "SMOTE for learning from imbalanced data: Progress and challenges, marking the 15-year anniversary," *J. Artif. Intell. Res.*, vol. 61, pp. 863–905, Apr. 2018, doi: 10.1613/jair.1.11192.

[11] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu, "Feature selection: A data perspective," *ACM Comput. Surv.*, vol. 50, no. 6, pp. 1–45, Nov. 2017, doi: 10.1145/3136625.

[12] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 3146–3154.

[13] E. Bjornson, J. Born, and A. Mardinoglu, "Personalized cardiovascular disease prediction and treatmenta review of existing strategies and novel systems medicine tools," *Frontiers Physiol.*, vol. 11, p. 705, Jan. 2020, doi: 10.3389/fphys.2020.00705.

[14] M. Cikes and S. D. Solomon, "Beyond ejection fraction: An integrative approach for assessment of cardiac structure and function in heart failure," *Eur. Heart J.*, vol. 42, no. 6, pp. 657–670, 2021, doi: 10.1093/eurheartj/ehaa731.

[15] P. Ponikowski et al., "2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure," *Eur. Heart J.*, vol. 37, no. 27, pp. 2129–2200, Jul. 2016, doi: 10.1093/eurheartj/ehw128.

[16] B. A. Goldstein, A. M. Navar, and R. E. Carter, "Moving beyond regression techniques in cardiovascular risk prediction: Applying machine learning to address analytic challenges," *Eur. Heart J.*, vol. 38, no. 23, pp. 1805–1814, 2017, doi: 10.1093/eurheartj/ehw302.

[17] A. C. Alba, T. Agoritsas, and M. Walsh, "Discrimination and calibration of clinical prediction models: Users' guides to the medical literature," *JAMA*, vol. 310, no. 19, pp. 2501–2502, 2013, doi: 10.1001/jama.2017.12126.

[18] S. E. Awan, F. Sohel, F. M. Sanfilippo, M. Bennamoun, and G. Dwivedi, "Machine learning in heart failure: Ready for prime time," *Current Opinion Cardiol.*, vol. 33, no. 2, pp. 190–195, 2018, doi: 10.1097/HCO.0000000000000507.

[19] M. Panahiazar, V. Taslimitehrani, N. L. Pereira, and J. Pathak, "Using EHRs for heart failure therapy recommendation using multidimensional patient similarity analytics," *Stud. Health Technol. Informat.*, vol. 210, p. 369, Jun. 2015, doi: 10.3233/978-1-61499-512-8-369.

[20] B. J. Mortazavi, N. S. Downing, E. M. Bucholz, K. Dharmarajan, A. Manhapra, S.-X. Li, S. N. Negahban, and H. M. Krumholz, "Analysis of machine learning techniques for heart failure readmissions," *Circulat., Cardiovascular Quality Outcomes*, vol. 9, no. 6, pp. 629–640, Nov. 2016, doi: 10.1161/CIRCOUTCOMES.116.003039.

[21] A. Bagheri, Q. Gao, and P. Rafi, "Genetic algorithm-based heart failure prediction and diagnosis system," *J. Intell. & Fuzzy Syst.*, vol. 37, no. 6, pp. 7649–7660, 2019, doi: 10.3233/JIFS-179230.

[22] J.-J. Beunza, E. Puertas, E. García-Ovejero, G. Villalba, E. Condes, G. Koleva, C. Hurtado, and M. F. Landecho, "Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease)," *J. Biomed. Informat.*, vol. 97, Sep. 2019, Art. no. 103257, doi: 10.1016/j.jbi.2019.103257.

[23] A. M. Alaa, T. Bolton, E. Di Angelantonio, J. H. F. Rudd, and M. van der Schaar, "Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 U.K. biobank participants," *PLoS ONE*, vol. 14, no. 5, May 2019, Art. no. e0213653, doi: 10.1371/journal.pone.0213653.

[24] E. Choi, A. Schuetz, W. F. Stewart, and J. Sun, "Using recurrent neural network models for early detection of heart failure onset," *J. Amer. Med. Inform. Assoc.*, vol. 24, no. 2, pp. 361–370, 2017, doi: 10.1093/jamia/ocw179.

[25] F. M. Zahid, S. Faisal, S. Ramzan, and I. Hussain, "A novel heart disease prediction method based on machine learning techniques and feature selection," *Healthcare Informat. Res.*, vol. 25, no. 2, pp. 124–135, 2019, doi: 10.4258/hir.2019.25.2.124.

[26] G. Guidi, M. C. Pettenati, P. Melillo, and E. Iadanza, "A machine learning system to improve heart failure patient assistance," *IEEE J. Biomed. Health Informat.*, vol. 18, no. 6, pp. 1750–1756, Nov. 2014, doi: 10.1109/JBHI.2014.2341731.

[27] N. Van Thieu and S. Mirjalili, "MEALPY: An open-source library for latest meta-heuristic algorithms in Python," *J. Syst. Archit.*, vol. 139, Jun. 2023, Art. no. 102871, doi: 10.1016/j.sysarc.2023.102871.

[28] Z. Qin, F. Yu, Z. Shi, and Y. Wang, "Adaptive inertia weight particle swarm optimization," in *Artificial Intelligence and Soft Computing–ICAISC* (Lecture Notes in Computer Science), vol. 4029, L. Rutkowski, R. Tadeusiewicz, L. A. Zadeh, and J. M. urada, Eds., Berlin, Germany: Springer, 2006, pp. 450–459, doi: 10.1007/11785231_48.

[29] R. Karthiga, G. Usha, N. Raju, and K. Narasimhan, "Transfer learning based breast cancer classification using one-hot encoding technique," in *Proc. Int. Conf. Artif. Intell. Smart Syst. (ICAIS)*, Mar. 2021, pp. 115–120, doi: 10.1109/ICAIS50930.2021.9395930.

[30] W. Satriaji and R. Kusumaningrum, "Effect of synthetic minority oversampling technique (SMOTE), feature representation, and classification algorithm on imbalanced sentiment analysis," in *Proc. 2nd Int. Conf. Informat. Comput. Sci. (ICICoS)*, Oct. 2018, pp. 1–5, doi: 10.1109/ICICOS.2018.8621648.

[31] R. Blagus and L. Lusa, "Joint use of over- and under-sampling techniques and cross-validation for the development and assessment of prediction models," *BMC Bioinf.*, vol. 16, no. 1, pp. 1–10, Dec. 2015, doi: 10.1186/s12859-015-0784-9.

[32] N. V. Chawla, "Data mining for imbalanced datasets: An overview," in *Data Mining and Knowledge Discovery Handbook*. Cham, Switzerland: Springer, 2009, pp. 875–886, doi: 10.1007/978-0-387-09823-4_45.

[33] A. Ishaq, S. Sadiq, M. Umer, S. Ullah, S. Mirjalili, V. Rupapara, and M. Nappi, "Improving the prediction of heart failure patients' survival using SMOTE and effective data mining techniques," *IEEE Access*, vol. 9, pp. 39707–39716, 2021, doi: 10.1109/ACCESS.2021.3064084.

[34] S. M. Alhashmi, M. S. I. Polash, A. Haque, F. Rabbe, S. Hossen, N. Faruqui, I. Abaker Targio Hashem, and N. Fathima Abubacker, "Survival analysis of thyroid cancer patients using machine learning algorithms," *IEEE Access*, vol. 12, pp. 61978–61990, 2024, doi: 10.1109/ACCESS.2024.3392275.

[35] S. D. Yang, Z. A. Ali, and B. M. Wong, "Fluid-gpt (fast learning to understand and investigate dynamics with a generative pre-trained transformer): Efficient predictions of particle trajectories and erosion," *Ind. & Eng. Chem. Res.*, vol. 62, no. 37, pp. 15278–15289, 2023.

[36] S. Tang and J. Wiens, "Model selection for offline reinforcement learning: Practical considerations for healthcare settings," in *Proc. Mach. Learn. Healthcare Conf.*, 2021, pp. 2–35.

[37] M. Ahmed, M. H. Sulaiman, M. M. Hassan, M. A. Rahaman, and M. Abdullah, "Selective opposition based constrained barnacle mating optimization: Theory and applications," *Results Control Optim.*, vol. 17, Dec. 2024, Art. no. 100487.

[38] M. Ahmed, M. H. Sulaiman, A. J. Mohamad, and M. Rahman, ''Gooseneck barnacle optimization algorithm: A novel nature inspired optimization theory and application,'' *Math. Comput. Simul.*, vol. 218, pp. 248–265, Apr. 2024.

MARZIA AHMED received the Ph.D. degree from the Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA), in 2024. She is an Assistant Professor with the Department of Software Engineering, Daffodil International University, Dhaka, Bangladesh. She is an inventor of the Gooseneck Barnacle Optimization Algorithm (GBOA), which has been applied in various optimization problems. She has authored more than 20 research papers published in international journals and conferences. Her research focuses on swarm intelligence, optimization algorithms, artificial intelligence, and the Internet of Things (IoT).

MOHD HERWAN SULAIMAN (Senior Member, IEEE) received the B.Eng. degree (Hons.) in electrical-electronics, the M.Eng. degree in electrical-power, and the Ph.D. degree in electrical engineering from the Universiti Teknologi Malaysia (UTM), in 2002, 2007, and 2012, respectively. He is an Associate Professor with the Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA). He has authored or co-authored more than 190 technical papers at the international journals and conferences. His research interests include power system optimization and swarm intelligence applications to power system and energy studies. He was invited as a journal reviewer and an associate editor for several international impact journals in the field of power system, soft computing application, and many more. He appeared in the world's top 2% of scientists according to Stanford University, in 2021, 2022, and 2023.

MD MARUF HASSAN (Member, IEEE) was born in Dhaka, Bangladesh. He received the bachelor's degree in information systems from Australian Catholic University, Australia, the master's degree in computer science and engineering from East West University, Bangladesh, and the Ph.D. degree in computer engineering from the Universiti Malaysia Perlis (UniMAP), Malaysia. He has amassed over 17 years of experience in both academic and professional roles, primarily in Australia and Bangladesh. He holds multiple industry certifications, including a Certified Information Systems Auditor (CISA), a Computer Hacking Forensic Investigator (CHFI), and a Certified Ethical Hacker (CEH). Currently, he is an Associate Professor with Southeast University, Bangladesh. Prior to this role, he was an Associate Professor with Daffodil International University, Bangladesh, where he was the Director of the Cyber Security Centre and coordinated cybersecurity programs. He has designed the curriculum and syllabus for several courses related to cybersecurity and has played a key role in developing the M.Sc. degree in cyber security syllabus for multiple universities. He has authored more than 48 research articles in various internationally indexed journals, book chapters, and conference proceedings. He is a member of several professional societies and has been actively involved in conducting workshops, seminars, and training sessions in the field of cybersecurity.

TOUHID BHUIYAN received the Ph.D. degree in information security from Queensland University of Technology (QUT), Australia, focusing on trust management for intelligent recommendations. He is a leading Expert in cyber security and a Professor with the School of IT, Washington University of Science and Technology (WUST), VA, USA. He is a Certified Ethical Hacker with a Cyber Security: Cyber Risk and Resilience Certificate from the University of Oxford. His academic career includes tenure with Monash University, The University of Western Australia, QUT, the University of Western Sydney, the University of Malaysia Perlis, and East West University. Previously, he held the position of the Director of the Cyber Security Centre. He led the Software Engineering Department and the Computer Science and Engineering Department, Daffodil International University (DIU), Bangladesh. He has authored over 150 research publications in prestigious journals, books, and conference proceedings. His research spans various areas including cyber security, intelligent recommendations, social networks, trust management, big data analytics, e-health, and e-learning. He received the Australian Postgraduate Award (APA) and the Deputy Vice-Chancellor's Initiative Scholarship from QUT.