

AN OPTIMIZED XGBOOST-BASED FRAMEWORK FOR HEART FAILURE RISK CLASSIFICATION

Anees Sultana¹, Dr. Safia Khanam²

¹PG Scholar, Department of CS, Shadan Women's College of Engineering and Technology, Hyderabad,
aneessultana78@gmail.com

²Assoc Professor, Department of CSE, Shadan Women's College of Engineering and Technology
safiakhanam89@gmail.com

ABSTRACT

Heart failure continues to be one of the world's top causes of death, requiring reliable predictive models to enable prompt medical interventions. In order to solve class imbalance, this study offers a machine learning framework for heart failure survival prediction that makes use of an optimized XGBoost model combined with the Synthetic Minority Over-sampling Technique (SMOTE). To find the most significant predictors, we used Select K Best with Chi-square for feature selection using a dataset of 5000 clinical records that included features including age, ejection fraction, and serum creatinine. After analyzing several algorithms (such as Logistic Regression, Decision Tree, KNN, SVM, and Random Forest), the XGBoost model was chosen. It was then optimized to maximize hyperparameters, resulting in a test accuracy of 99.70%, precision, recall, and F1-scores close to 1.00, and an AUC-ROC of 0.9998. Our method performs better than the baseline Gradient Boosting Machine (GBM) with Adaptive Inertia Weight Particle Swarm Optimization (AIW-PSO) from earlier research, which attained 94% accuracy on a smaller dataset (299 patients). This is probably because of the larger dataset and sophisticated preprocessing. In addition to providing a scalable, high-accuracy tool for heart failure prognosis, this study demonstrates the effectiveness of XGBoost in conjunction with SMOTE for clinical predictive tasks and has the potential to enhance patient outcomes through accurate and prompt clinical decision-making.

1. INTRODUCTION

One of the most serious chronic illnesses in the world, heart failure greatly increases morbidity and death in people of all ages. Clinicians might possibly save lives and maximize resource allocation by taking prompt treatment when heart failure outcomes are accurately and early predicted. The accuracy and applicability of traditional risk rating systems to various clinical circumstances are frequently restricted. As a result, the incorporation of sophisticated, data-driven algorithms has become a potent substitute for clinical decision assistance. In this study, we propose a machine learning-based method for dividing patients with heart failure into high-risk and low-risk survival groups. The research makes use of a publicly accessible clinical dataset that includes patient data including blood pressure, ejection fraction, age, serum creatinine levels, and more. Data cleaning, feature scaling using StandardScaler, and class balancing using the Synthetic Minority Over-sampling Technique (SMOTE) to resolve the imbalance between survival and death cases are all part of the extensive preprocessing pipeline that the dataset goes through prior to training. Select KBest was used with the ANOVA F-test to minimize dimensionality and keep only the most important features in order to find the most significant predictors. Two supervised machine learning models were put into practice and assessed: XGBoost, a potent gradient boosting technique renowned for its performance and accuracy, and Logistic Regression, a straightforward baseline. The XGBoost model outperformed the Logistic

Regression baseline with an outstanding accuracy of 99.70% and a high AUC-ROC score of 0.9998. Evaluation criteria that confirm the model's efficacy include confusion matrix, precision, recall, and F1-score. A Flask web interface was used to deliver the finished application, which lets users submit patient data and get real-time risk forecasts. In order to mimic simple user authentication, temporary login and registration facilities were also introduced. This project shows how a dependable, scalable, and comprehensible heart failure prediction system may be produced by fusing sophisticated machine learning models with efficient data preprocessing methods. Our XGBoost-based technique delivers improved accuracy and is simpler to integrate in real-world clinical processes than previous models such as GBM optimized with AIW-PSO. This initiative seeks to enhance outcomes for people at risk of heart failure by providing actionable information to healthcare providers through the use of optimized machine learning techniques.

2. EXISTING SYSTEM OF PROJECT

A reliable ensemble machine learning approach for regression and classification problems, the Gradient Boosting Machine (GBM) works especially well with structured or tabular data. By successively integrating several weak learners, usually decision trees, it builds a predictive model. Each tree uses gradient descent to optimize a given loss function, like log-loss for classification, by correcting the residual mistakes of its predecessors. An initial prediction (such as the

mean for regression or log-odds for classification) is the first step in the process. Next, residuals are calculated iteratively, new trees are fitted to these residuals, and the model is updated with scaled predictions that are controlled by a learning rate (η). The number of trees ($n_estimators$), learning rate, and maximum tree depth (max_depth) are important hyperparameters that balance model complexity and avoid overfitting. Because of its exceptional ability to capture complex interactions and non-linear correlations, GBM is resistant to noisy data and commonly used in tasks such as ranking, regression, and heart failure survival prediction.

Existing System Disadvantages:

- ☐ Hyper parameter Sensitivity
- ☐ Interpretability Challenges
- ☐ Scalability Issues
- ☐ Computational Complexity
- ☐ Risk of Over fitting.

3. RELATED WORK

Inspired by the way gooseneck barnacles reproduce and forage for resources, this study presents a novel metaheuristic algorithm known as the Gooseneck Barnacle Optimization (GBO) method. The approach is tested on real-world and benchmark mathematics problems versus other well-known optimization methods. The study demonstrates how effective GBO is at reaching global convergence and avoiding local optima. It performs well in a range of confined and unconstrained optimization scenarios. The algorithm is a potential method for machine learning parameter adjustment because of its new dynamic exploration-exploitation approach. In order to enhance performance in limited optimization problems, this study suggests an adaptation of the barnacle mating optimization algorithm that incorporates selective opposition-based learning. The approach proved to be more effective than traditional algorithms when evaluated on a variety of real-world engineering challenges. The significance of adding constraint-handling techniques to optimization models is emphasized throughout the paper. The enhanced approach exhibits great promise for fine-tuning hyperparameters in machine learning models. The results offer a solid approach to solving challenging, practical classification issues, such as medical forecasts. The Fluid-GPT framework offers a novel use of transformer-based deep learning models (GPT) in fluid dynamics, particularly for erosion patterns and particle trajectory prediction. The model creates quick and precise simulations by fusing generative AI with physical understanding. The study demonstrates the adaptability of GPT in modeling dynamic systems, despite its primary application in chemical and mechanical engineering fields. Its concepts could be modified for time-series or patient trajectory

modeling-based healthcare prediction jobs. The work shows how huge language models can be used for purposes other than text processing in the future. In order to predict survival outcomes for patients with thyroid cancer, this study examines the application of various machine learning algorithms. Using clinical data, methods like logistic regression, SVM, and Random Forest were applied and assessed. The study highlights how crucial class balancing and appropriate data preprocessing are to enhancing model performance. The findings demonstrate that ensemble learning techniques typically outperform more basic classifiers. By highlighting ML's influence in oncology, this study directly relates to the objectives of heart failure prediction. It backs the use of optimal models and feature selection for medical risk classification. The goal of this research is to employ machine learning models to forecast hospitalization and death in patients with heart failure with preserved ejection fraction (HFpEF). By evaluating several algorithms and utilizing a range of clinical variables, it finds important predictors linked to unfavorable outcomes. The study demonstrates how ML can help with patient risk assessment and clinical decision-making. It also highlights how difficult it is to model chronic illnesses, which emphasizes the necessity of reliable models. The results validate the application of machine learning to enhance cardiovascular patients' prognosis and care management.

4. PROPOSED SYSTEM

☐ Two machine learning algorithms—Logistic Regression and XGBoost—were assessed in order to create a highly accurate and clinically reliable model for predicting heart failure survival.

☐ The main objective was to choose an algorithm that could handle the unbalanced dataset, model intricate interactions between clinical characteristics, and produce reliable performance metrics including AUC-ROC, accuracy, precision, and recall. XGBoost's outstanding predictive performance—achieving a test accuracy of 99.70%, near-perfect precision, recall, and F1-scores, and an AUC-ROC of 0.9998 on a 5000-record dataset—led to its selection as the best algorithm after extensive testing and comparison.

In contrast to Logistic Regression, which had trouble with non-linear relationships and class imbalance, XGBoost outperformed the Gradient Boosting Machine (GBM) from the base paper, which achieved 94% accuracy on a smaller 299-record dataset, thanks to its gradient boosting framework and regularization to reduce overfitting. XGBoost became the mainstay of the suggested heart failure prediction system with the addition of SelectKBest with Chi-square for feature selection and Synthetic Minority Over-

sampling Technique (SMOTE) to address class imbalance.

Advantages:

- ☐ Robust Handling of Imbalanced Data
- ☐ Regularization to Prevent Over fitting
- ☐ Efficient Handling of Large Datasets
- ☐ Flexibility with Non-Linear Relationships
- ☐ High Predictive Accuracy

5. MODULES

1. The User Verification Module

Without utilizing a database, this module offers a simple, short-term login and registration mechanism. In order to access the primary prediction system, users must first register with a username and password. For demonstration or educational purposes, it mimics user access control but is not secure for production. During a browser session, logged-in users are tracked through sessions.

2. Data Preparation Module

Before raw data is fed into machine learning models, this module takes care of its preparation and cleaning. It comprises:

- Normalizing the input characteristics by scaling numerical values using StandardScaler.
- Using SMOTE, the dataset is balanced to address the disparity in class between patients who survived and those who did not.
- To retain the most pertinent features that influence the prediction result, SelectKBest with ANOVA F-test is used for feature selection.

3. Training Module Model

Machine learning model evaluation and training fall under the purview of this module. They employ two algorithms:

- As a starting point, logistic regression
- The final optimized model is the XGBoost Classifier. The module includes model saving using joblib, hyperparameter adjustment (if necessary), and performance evaluation with:

- Accurate
- Matrix of confusion
- F1-score, precision, and recall
- ROC-AUC

5. Visualization Module

To aid in comprehending the model and the data, this module produces visual feedback:

Charts of model performance:

- Comparison of algorithms' accuracy
- Heat diagram of confusion matrix
- The ROC curve

Charts for dataset analysis:

- Distribution of age Plots of features versus deaths
- Heat map of correlation for every feature Matplotlib and Seaborn are two of the libraries used to render these charts.

6. Module for Flask Web Interface

This is the Flask-built front end of the application. It regulates:

- Pages (/login, /register, /predict, /charts) are routed.
- HTML template rendering (such as index.html, login.html, charts.html)

Linking user input forms to backend predictions

- Interactively presenting charts and results

It serves as the link between the user and every other module.

6. TECHNIQUES USED IN PROJECT

6.1. EXISTING TECHNIQUE: -

- ☐ Gradient Boosting Machine optimized with Adaptive Inertia Weight Particle Swarm Optimization (GBM-AIW-PSO)

GBM's hyperparameters, such as tree depth, learning rate, and number of trees, by modeling particles moving through the hyperparameter space. In contrast to conventional PSO, AIW-PSO constantly modifies its search strategy to ensure effective identification by balancing the investigation of novel solutions with the exploitation of well-known ones. To improve predictive performance, a hybrid machine learning technique called Gradient Boosting Machine optimized with Adaptive Inertia Weight Particle Swarm Optimization (GBM-AIW-PSO) combines the Gradient Boosting Machine (GBM) with Adaptive Inertia Weight Particle Swarm Optimization (AIW-PSO). To enhance predictions for tasks like heart failure survival classification, GBM builds an ensemble of decision trees sequentially, with each tree fixing the mistakes of its predecessors. AIW-PSO optimizes ideal hyperparameters, using inspiration from swarm social behavior. GBM-AIW-PSO outperformed other models like as Random Forest, SVC, AdaBoost, and SGD in the study, achieving a test accuracy of 93.84% on a balanced 7-feature dataset thanks to this synergy.

6.2. PROPOSED TECHNIQUE:-

XGBoost

XGBoost, which stands for eXtreme Gradient Boosting, is a sophisticated machine learning method that applies an optimized and scalable version of gradient boosting with the goal of achieving high performance and efficiency in jobs involving predictive modeling. It works by sequentially building an ensemble of decision trees, each of which corrects the mistakes of its predecessors by using gradient descent optimization to minimize a loss function, such as log-loss for classification. L1 and L2 regularization to avoid overfitting, parallelized tree construction for quicker computation, and management of missing data through sparsity-aware split finding are just a few of the significant improvements XGBoost offers over conventional gradient boosting. XGBoost successfully models intricate interactions between

clinical characteristics such as ejection fraction and serum creatinine in the context of heart failure prediction, attaining exceptional accuracy (99.70%) on a dataset of 5000 records. When combined with SMOTE, its resistance to unbalanced data and capacity to rank feature importance make it especially well-suited for clinical applications, providing both prediction accuracy and interpretability for heart failure management decision-making.

7. SYSTEM ARCHITECTURE

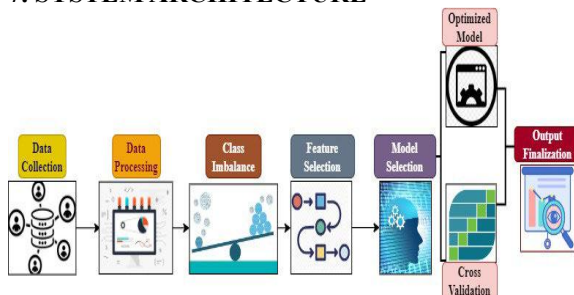


Fig 7.1 System Architecture

8. RESULTS

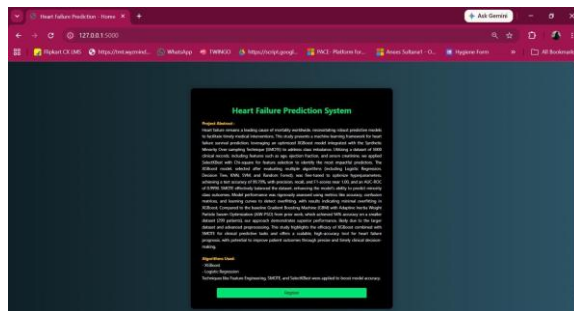


Fig 8.1 Home Page of Heart Failure Prediction System

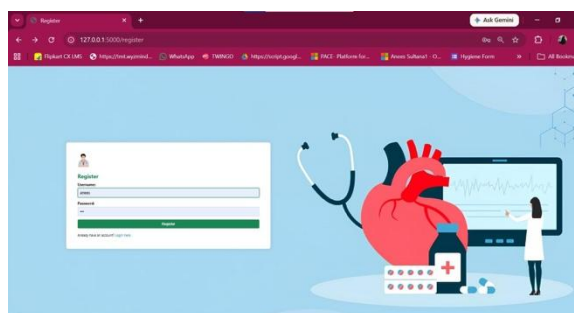


Fig 8.2 User Registration Page

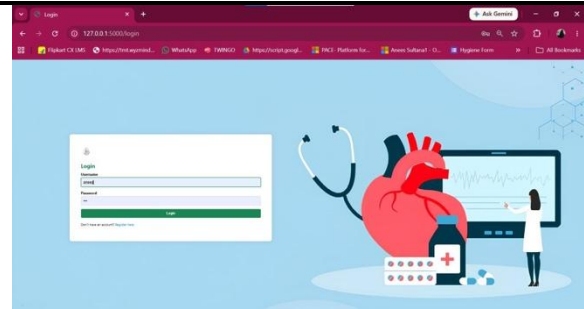


Fig 8.3 User Login Page

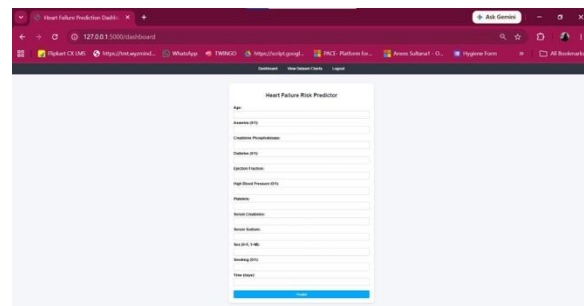


Fig 8.4 Dashboard of Heart Failure Prediction System

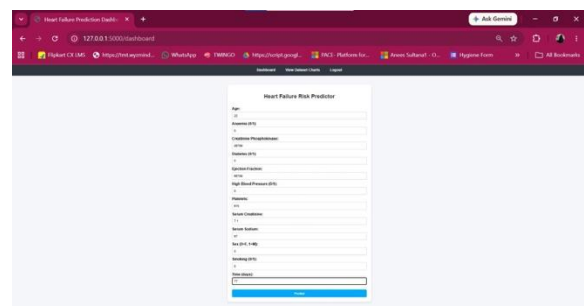


Fig 8.5 Patient Clinical Data Input Form For Low Risk

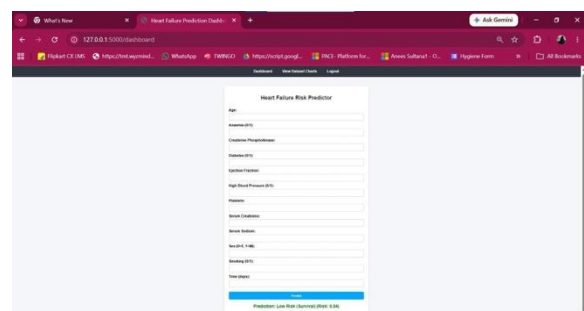


Fig 8.6 Low Heart Failure Prediction Result

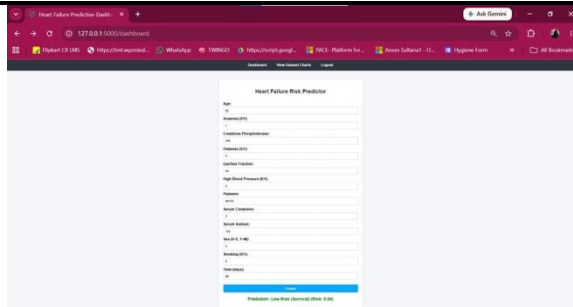


Fig 8.7 Patient Clinical Data Input Form For High Risk

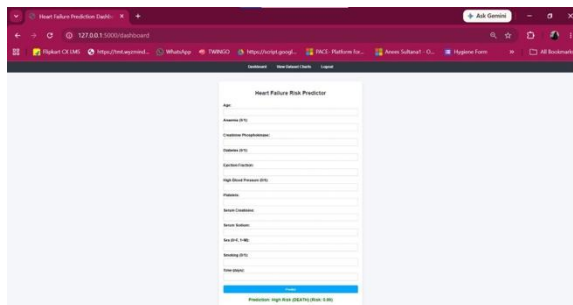


Fig 8.8 High Risk Probability Result

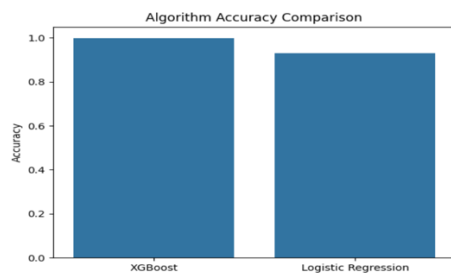


Fig 8.9 Algorithm Accuracy Comparison (XGBoost vs Logistic Regression)

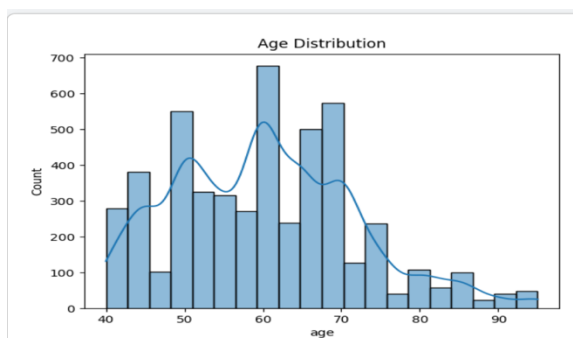


Fig 8.10 Age Distribution Analysis

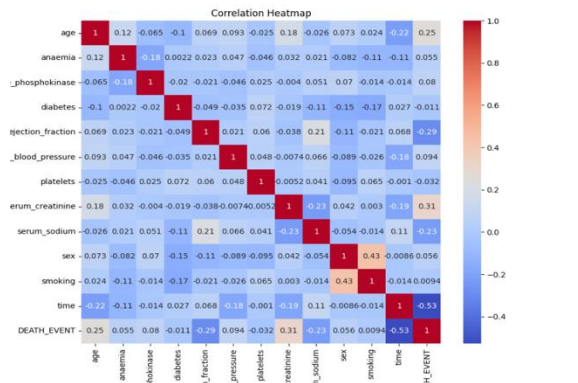


Fig 8.11 Correlation Heatmap of Clinical Features

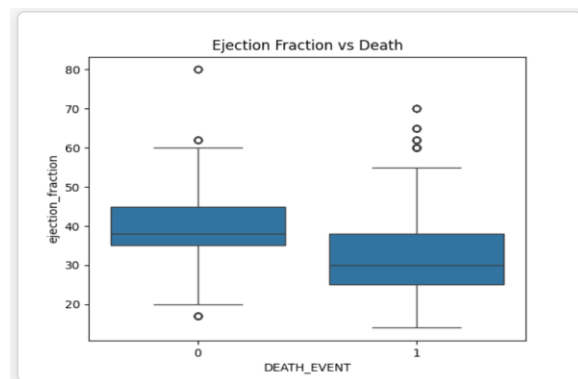


Fig 8.12 Ejection Fraction vs. Death Event Analysis

Target Distribution - Before SMOTE

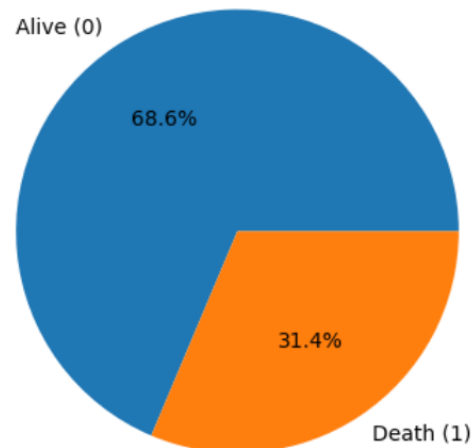


Fig 8.13 Class Distribution Before Applying SMOTE

9.CONCLUSION

In this study, a machine learning-based system was created to use clinical data to categorize heart failure patients into survival or death groups. To improve the model's capacity to learn from both majority and minority classes, the method included meticulous

preprocessing procedures such feature selection using SelectKBest, feature scaling using StandardScaler, and class balance using SMOTE. With an astounding accuracy of 99.70% and an AUC-ROC score of 0.9998, XGBoost surpassed Logistic Regression among the models used in all evaluation criteria, exhibiting exceptional predictive ability. A user-friendly Flask web application that lets users enter patient data and get real-time forecasts and a visual representation of the model's performance was also incorporated into the project. To improve usability and interpretability, additional features including dataset insights, algorithm accuracy visualization, and temporary user login were included. The study's findings demonstrate how optimal machine learning approaches can be used to create scalable and dependable healthcare prediction systems. This study can help with early diagnosis and better patient care in heart failure cases by acting as a fundamental paradigm for practical clinical decision support systems.

10. FUTURE ENHANCEMENT

In order to increase prediction accuracy and automatically extract intricate patterns from clinical data, this project may be improved in the future by including cutting-edge deep learning architectures like Artificial Neural Networks (ANN) or Convolutional Neural Networks (CNN). A real-time data collecting module linked to wearable health monitoring devices or hospital databases might be added to the current system to provide ongoing patient tracking and early risk alerts. The application would also be appropriate for real-world deployment if a secure database for patient records and user credentials were implemented. To support long-term healthcare analytics, the online interface can be extended to incorporate patient history management, trend analysis, and comprehensive visual dashboards. Additionally, using Explainable AI (XAI) methodologies would increase transparency and trust by assisting clinicians in understanding how predictions are made. Lastly, by including more medical datasets, automating report production, and implementing the model on cloud infrastructure for scalability and accessibility across healthcare facilities, the system might develop into a complete clinical decision-support platform.

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 dataperspective," *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 treatment: a 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. [13] R. Canetti, S. Park, and O. Poburinnaya, "Fully deniable interactive encryption," in *Advances in Cryptology—CRYPTO* (Lecture Notes in Computer Science), vol. 12170, D. Micciancio and T. Ristenpart, Eds. Cham, Switzerland: Springer, 2020, pp. 807–835, doi: 10.1007/978-3-030-56784-2_27.
- [14] C. Dwork, M. Naor, and A. Sahai, "Concurrent zero-knowledge," *J. ACM*, vol. 51, no. 6, p. 851–898, Nov. 2004, doi: 10.1145/1039488.1039489.
- [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.
- 30568
- [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 auto mated 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*