Prediction Of Heart Failure Patient Survival Using Machine Learning

Merlin Alex[1], **Maheen Hasib**[2]
[1]Heriot-Watt Dubai University, Dubai, United Arab Emirates
[2]Heriot-Watt Dubai University, Dubai, United Arab Emirates
merlinalex2012@yahoo.in, M.Hasib@hw.ac.uk

Abstract

Heart failure is a common, incurable illness with high morbidity and mortality rates globally. Early detection enables proper medication and monitoring, significantly affecting survival rates especially in the present healthcare environment driven by the covid-19 pandemic. Machine learning techniques offer a powerful tool for predicting heart failure survival by uncovering hidden data patterns.

This paper aims to predict heart failure patient survival using machine learning, focusing on identifying significant features through techniques like Recursive feature elimination, Random Forest and Information Gain. The machine learning model used is an ensemble method, a stacked classifier with CatBoost, LightGBM and XGBoost as base models and Multilayer Perceptron as the meta-learner. The best-performing models were selected from a list of trained models, including Logistic Regression, Random Forest, Support Vector Machine, LightGBM, CatBoost, XGBoost, and Multilayer Perceptron. The heart failure patient survival dataset utilized in this research is artificially created through Gretel AI, a synthetic data generation platform, based on a primary UCI medical dataset (Heart Failure Clinical Records, UCI). This method ensures data confidentiality while preserving essential statistical characteristics. This study contributes to research on predicting heart failure survival by emphasizing early intervention and demonstrating the potential of machine learning in improving patient outcomes.

Introduction

Heart failure (HF) is a chronic condition where the heart fails to pump blood efficiently, leading to symptoms such as shortness of breath, fatigue, and swelling in the limbs and abdomen (NHLBI 2022). It is a progressive syndrome resulting from structural or functional cardiac abnormalities impairing the heart's ability to circulate blood adequately (Masetic and Subaşi 2016). Though HF itself is not classified as a disease, it significantly contributes to life-threatening cardiovascular conditions such as coronary artery disease, cardiomyopathy, and myocarditis (Quinn 2006). Over 64 million people globally suffer from HF, with its prevalence expected to increase by 46% between 2012 and 2030 (Virani et al. 2021; Kapelios et al. 2023). The rising burden of HF necessitates proactive healthcare strategies to reduce its social and financial impact (Kapelios et al. 2023; Savarese et al. 2022).

The European Society of Cardiology (ESC) 2016 guidelines define HF as a clinical syndrome characterized by breathlessness, ankle swelling, and fatigue, caused by cardiac abnormalities leading to reduced cardiac output and/or elevated intracardiac pressures at rest or during stress (Hu et al. 2022). However, the 2021 universal definition expands this by incorporating elevated natriuretic peptide levels or objective evidence of systemic congestion as diagnostic criteria (Hu et al. 2022). Heart failure is further classified into two main types based on ejection fraction:

- Heart Failure with Reduced Ejection Fraction (HFrEF): Left ventricular ejection fraction (LVEF)≤ 40%, commonly associated with impaired systolic function due to myocardial infarction or dilated cardiomyopathy.
- Heart Failure with Preserved Ejection Fraction (HFpEF): LVEF ≥ 50%, where the heart contracts normally but struggles to relax and fill properly, often linked to hypertension and diabetes (Inamdar and Inamdar 2016; Ponikowski et al. 2016).
- A third category, Heart Failure with Mid-Range Ejection Fraction (HFmrEF) (LVEF 41-49%), represents an intermediate state (Ponikowski et al. 2016). The New York Heart Association (NYHA) classification and ACC/AHA staging system are also widely used to assess HF severity based on symptoms and disease progression.

These classifications guide treatment strategies, where HFrEF is managed with beta-blockers and ACE inhibitors, while HFpEF treatment focuses on symptom management (Siriwardena and Fan 2018).

Traditional diagnostic methods for HF rely on clinical symptoms, echocardiography, biomarkers (BNP, NT-proBNP), and exercise capacity assessment (Ponikowski et al. 2016). However, these approaches often fail to provide early detection, leading to delayed interventions and poor

patient outcomes. Moreover, ethnic disparities affect HF prognosis, as seen in Bangladeshi and Pakistani populations, where genetic differences impact response to treatment (NIHR 2023). The dataset used in this study focuses on the Pakistani population, where 200,000 new cases of coronary heart disease are reported annually (Moreno-Sanchez 2023). The long-term effects of COVID-19 have further exacerbated cardiovascular complications, with studies showing a higher incidence of HF among recovered individuals (Zuin et al. 2022). Preventive healthcare measures such as vaccination may play a role in reducing these risks.

With advancements in medical imaging, biomarkers, and wearable technology, machine learning (ML) has emerged as a crucial tool for predicting HF outcomes. Traditional risk scores such as the Seattle Heart Failure Model (SHFM) and Heart Failure Survival Score (HFSS) integrate clinical and imaging parameters to assess survival probabilities (Pocock et al. 2012). However, ML models can analyze complex datasets, extract hidden patterns, and provide personalized risk predictions. This paper explores various ML techniques, including Logistic Regression, Random Forest, Support Vector Machine (SVM), LightGBM, CatBoost, and XGBoost, alongside an ensemble learning approach to enhance HF survival prediction. The study also integrates feature selection techniques such as Recursive Feature Elimination (RFE), Random Forest, and Information Gain to improve model interpretability and accuracy.

The evolution of HF diagnosis has been shaped by medical advancements. Early approaches relied on clinical characteristics and echocardiography, while later innovations introduced biomarkers (BNP, NT-proBNP), exercise capacity assessment, and risk scores for survival prediction (Ponikowski et al. 2016; Hunt et al. 2009; Lang et al. 2015). The discovery of X-rays, electrocardiography (ECG), and echocardiography revolutionized HF assessment, while modern AI-driven techniques enable real-time prognosis (Ribatti 2009; Davis, Hobbs, and Lip 2000; Hajar 2019). This research aims to bridge existing gaps in HF survival studies, such as the need for diverse datasets, robust model performance, and deeper understanding of feature relationships. The findings aim to enhance predictive analytics for HF survival and improve healthcare decision-making in resource-limited settings.

Related work

Heart failure is a progressive condition with varying symptoms, development rates, and causes, making survival prediction complex (Sabbah 2016). Even in clinically stable patients, the silent nature of HFrEF can increase mortality risk due to underdiagnosis or inadequate treatment. Additionally, patient-specific factors such as age, gender, genetics, lifestyle, and socioeconomic status contribute to prognosis variability, complicating survival predictions (Sciomer et al. 2020). Comorbidities such as diabetes, hypertension, and chronic kidney disease further complicate patient outcomes and treatment plans, adding to the heterogeneity of HF cases (Zheng et al. 2021). These challenges highlight the need for advanced predictive models capable of handling multi-faceted patient data.

Machine learning has revolutionized medical diagnosis by discovering hidden patterns in complex datasets, enabling personalized risk prediction, early diagnosis, and improved patient outcomes (Bourazana et al. 2024). Supervised learning techniques such as Support Vector Machines (SVM), Decision Trees (DT), Random Forest (RF), and XGBoost have been extensively studied for HF survival prediction. A study employing SVM with a Gaussian kernel on a dataset of 299 patients achieved an accuracy of 85.71%, outperforming Decision Trees (78.67%), Random Forest (82.67%), and K-Nearest Neighbours (76%) (Sang et al. 2020). However, the study did not employ feature selection or data balancing techniques, which could have enhanced the model's performance and generalizability. Another study utilized synthetic minority oversampling technique (SMOTE) and Random Forest for feature ranking, identifying Ejection Fraction, Serum Creatinine, and Age as the most influential features. Using Extra Trees Classifier, the model achieved 92.6% accuracy, outperforming Logistic Regression (84.2%), Decision Tree (87.7%), and XGBoost (85.33%) (Ishaq et al. 2021). However, this study did not explore ensemble or neural network approaches, potentially limiting the exploration of complex relationships between features.

Few studies have explored unsupervised learning for HF survival prediction. A notable work applied K-Means clustering to segment HF patients into groups, followed by SVM training on these clusters. By optimizing hyperparameters through grid search, the model achieved an accuracy of 93.33% when using six clusters (Saravanan and Swaminathan 2021). The study demonstrated that clustering before classification could improve prediction performance by reducing intra-class variability but however, this approach requires careful cluster validation and interpretability checks before clinical application.

Ensemble learning methods combine multiple models to enhance prediction accuracy. A study introduced Balanced Random Forest (BRF) to tackle class imbalance, yielding an accuracy of 76.25%, outperforming SVM (68.89%) and KNN (66.89%) (Newaz et al. 2021). Another comparative study found that LightGBM outperformed traditional classifiers, achieving 85% accuracy, while XGBoost scored 84% (Mamun et al. 2022). Additionally, an experiment using XGBoost with feature permutation and hyperparameter tuning recorded an accuracy of 90% highlighting that advanced feature selection and tuning can significantly enhance model performance. (Kaushik and Birok 2021).

Neural networks (ANNs), particularly Multi-Layer Perceptrons (MLP), have demonstrated strong predictive capabilities in HF analysis. In a heart disease detection study, MLP trained with Particle Swarm Optimization (PSO) outperformed other models, achieving an accuracy of 84.61% (Bataineh and Manacek 2022). Although deep learning methods capture intricate data relationships, they require larger datasets, extensive tuning, and higher computational resources for optimal performance (Sarra, Dinar, and Mohammed 2022).

One major limitation in HF survival prediction is reliance on small datasets (299 records), leading to potential biases and overfitting (Ahmad et al. 2017). Many studies handled class imbalance using SMOTE, suggesting it as the preferred resampling method (Joloudari et al. 2023). However, the use of unsupervised learning and neural networks remains underexplored, likely due to dataset constraints and clinical preference for interpretable models (Tufail et al. 2023). A summary of these recent studies and their reported model accuracies is provided in Table 1.

Previous Studies	Dataset Records	ML Models Used	Highest Accuracy (%)
Chicco and Jurman (2020)	299	LR, RF, One Rule, DT, ANN, SVM radial, SVM linear, Gradient Boosting, K-NN, Naïve Bayes.	74 (RF)
Sang et al. 2020)	299	SVM, K-NN, DT, RF, XGBoost	85.71 (SVM)
Ishaq et al. (2021)	299	DT, AdaBoost, GBM, RF, Extra Trees, LR, Gaussian Naïve Bayes, SVM	92.6 (Extra Trees)
Saravanan and Swaminathan (2021)	299	SVM, K-Means	93.33 (SVM)
Newaz et al. (2021)	299	BRF, SVM, K-NN, RF	76.25 (BRF)
Kaushik and Birok (2021)	299	XGBoost, KNN, DT, RF, Naïve Bayes, etc.	90 (XGBoost)
Mamun et al. (2022)	299	LightGBM, SVM, DT, LR, Bagging, XGBoost	85 (LightGBM)
Bataineh and Manacek (2022)	303	MLP-PSO, LR, SVM, DT, RF, K-NN	84.61 (MLP-PSO)
Li et al. (2022)	4540	DT, SVM, Naïve Bayes, MLP, KNN, XGBoost, AdaBoost, LightGBM, LR, CatBoost and Bagging.	AUC: 0.833 (CatBoost)

Table 1: Comparison of ML Models in Previous Studies

This research builds upon previous work by integrating ensemble learning and feature selection techniques to improve HF survival prediction while addressing the challenges posed by small datasets and class imbalance.

Methodology

Phases of Implementation. The methodology of this study is organized into structured phases, each addressing critical steps in the prediction modelling process. The phases as shown in Figure 1 include dataset generation, preprocessing, exploratory data analysis (EDA), feature selection, model training, and model evaluation, clearly defining these phases facilitates reproducibility and clarity of the research process.

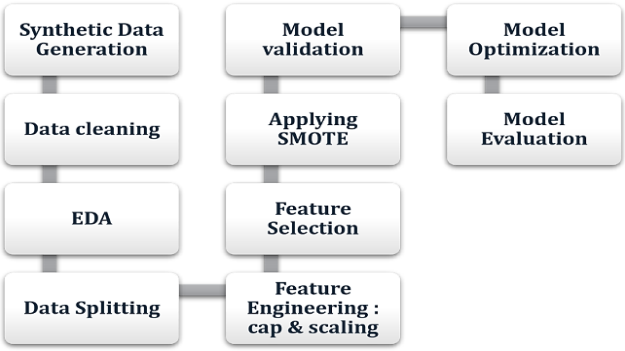

Figure 1: Phases of Implementation

Synthetic Dataset Generation. This study explored two synthetic data generation platforms, Mostly AI and Gretel AI, to create a synthetic dataset based on the UCI Heart Failure Clinical Records dataset (Ahmad et al. 2017). The original UCI dataset contains 299 patient records collected from the Institute of Cardiology and Allied Hospital in Faisalabad, Pakistan. However, this dataset was not directly used in the study; instead, it served as a reference for generating a synthetic dataset that maintained similar statistical properties.

Mostly AI generated a synthetic dataset with a 78.4% similarity accuracy to the original dataset, while Gretel AI, using its Navigator fine-tuning model, produced a synthetic dataset with a higher similarity accuracy of 91%. Due to this higher similarity, Gretel AI was selected to generate the synthetic dataset for this study. After multiple iterations and ensuring that no duplicate rows were retained, the final dataset comprised 1,072 records with 13 features, effectively replicating the structure and distribution of the original dataset without using any real patient data. However, it is essential to acknowledge that synthetic data, despite high statistical similarity, may not fully capture nuanced clinical variations present in real-world scenarios.

Dataset Preprocessing. Preprocessing involved handling missing values, normalizing features, and detecting outliers using statistical methods. These steps ensured data quality and suitability for modelling. Exploratory data analysis (EDA) was conducted to understand data distributions, identify outliers, and visualize feature relationships using techniques such as histograms, boxplots, scatter plots, stacked bar plots, pie charts, and correlation heatmaps. Visualizations such as correlation heatmaps provide valuable insights into feature interactions and inform subsequent feature selection. Table 2 provides an overview of the synthetic dataset, detailing the description and characteristics of each feature included in the study.

Features	Description	Outcomes
Age	Age of the patient	Range: 40 – 95 *(Years)*
Anaemia	Deficiency of haemoglobin or red blood cells in the blood.	0= Absence 1= Presence
Creatinine Phosphokinase	Level of Creatinine Phosphokinase (CPK) enzyme in the blood.	Range: 30 - 7702 *(mcg/L)*
Diabetes	Elevated levels of blood sugar in the body	0= Absence 1= Presence
Ejection Fraction	Percentage of blood leaving the heart each time it squeezes.	Range: 14-80
High Blood Pressure	when pressure in blood vessels is too high.	0= Absence 1= Presence
Platelets	Small blood cells that help with clotting to stop bleeding.	Range: 25.1 – 850 (*kiloplatelets /ML)*
Serum Creatinine	Level of creatinine in blood.	Range: 0.6 – 9.0 *(mg/dL)*
Serum Sodium	Level of sodium in blood.	Range: 125 - 148 *(mEq/L)*
Sex	Patient's gender.	0= Female 1= Male
Smoking	Patient's smoking habit.	0= Non-Smoker 1= Smoker
Time	Follow-up period.	Range: 4- 285 *(Days)*
Death Event (Target)	Occurrence of death during follow-up period.	0= Survived 1= Deceased

Table 2 : Synthetic Dataset Description

Exploratory Data Analysis (EDA). EDA techniques revealed important data distributions and relationships among features. The correlation heatmap, for instance, highlighted the relationship between different features and identified strong negative correlation between DEATH_EVENT and Time , as shown in Figure 2. Recognizing correlations is crucial, as highly correlated features could lead to redundancy and impact model interpretability.

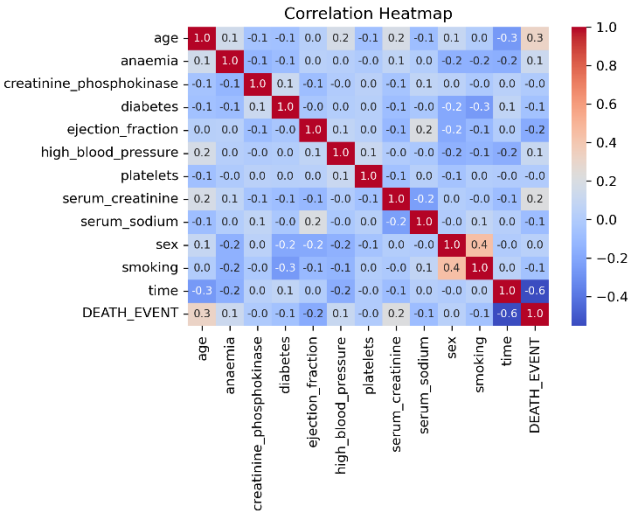

Figure 2: Synthetic Dataset Heatmap

Feature Engineering. During data preprocessing, outliers were identified in several features. To address this, capping at 0.05 and 0.95 percentiles was used, as it outperformed more extreme thresholds. StandardScaler was then used to normalize features, ensuring consistent scaling across bi-

nary and continuous variables. These steps were done separately for training and testing sets to avoid data leakage. Although SMOTE was considered here, it was applied only after feature selection to prevent biasing feature importance. This ensured transparent and unbiased methodology, with further details in the Model Training section.

Feature Selection. Feature selection techniques were employed to enhance model performance and interpretability. Recursive Feature Elimination (RFE) identified the optimal number of features through cross-validation, as illustrated in Figure 3, which determined that four features provided the best performance balance. This method systematically removed less significant features to improve the model's predictive accuracy while reducing complexity. Random Forest and Information Gain techniques also validated these results, collectively identifying Age, Ejection Fraction, Serum Creatinine, and Time as the most influential features. Random Forest assigned importance scores based on how much each feature contributed to reducing impurity in decision trees, while Information Gain quantified the relevance of individual attributes in classification performance. The selected features demonstrated clinical significance, aligning with medical studies indicating that ejection fraction and serum creatinine levels are strong indicators of heart failure progression. Additionally, time (follow-up period) played a crucial role, highlighting the impact of monitoring duration on survival outcomes. To ensure robustness, different feature selection thresholds were tested, including retaining six or three features, but the best performance was consistently achieved with four features, as identified by RFE and Random Forest.

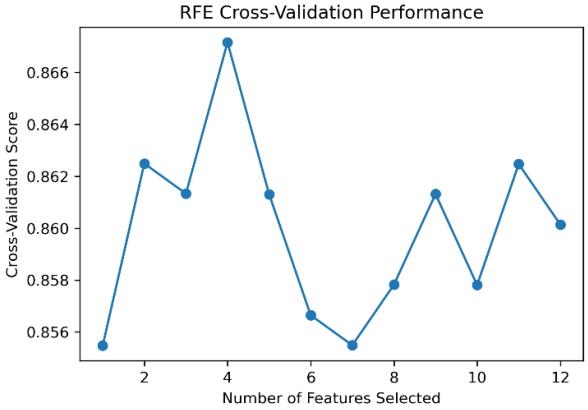

Figure 3: Optimal Number of Features Graph

Addressing Data Imbalance. Before training the machine learning models, SMOTE was applied to address class imbalance in the dataset. Since the original dataset had an imbalanced distribution of survival outcomes, applying SMOTE ensured a balanced class representation. This prevented models from being biased toward the majority class, improving recall and F1-score for minority class predictions. The impact of SMOTE is illustrated in Figure 4, showing the class distribution before and after resampling.

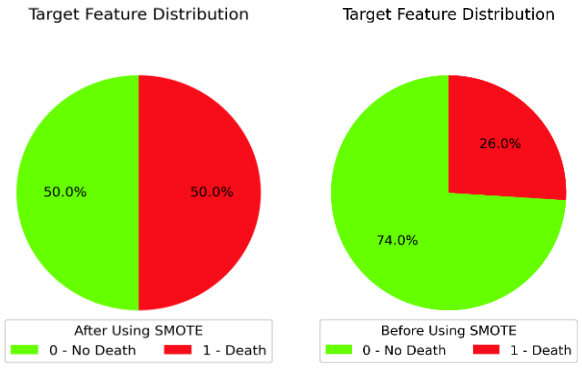

Figure 4: Target Feature Pie Charts Before and After SMOTE

Model Training. Several machine learning models were implemented, including Logistic Regression, Random Forest, Support Vector Machine (SVM), LightGBM, CatBoost, XGBoost, and Multilayer Perceptron (MLP). The models were trained and evaluated under three distinct testing scenarios:

- **Test 1**: All models were trained and evaluated without feature selection, using all 12 features.
- **Test 2**: Models were trained and evaluated using only the selected features identified from the feature selection process.
- **Test 3**: Based on the results from Test 1, the top-performing models—CatBoost, LightGBM, and XGBoost—were chosen as base models, with MLP as the meta-learner to develop a stacked ensemble model.

Despite minor accuracy reductions in Test 2, feature selection was retained due to its advantages in interpretability, reduced computational complexity, and mitigation of overfitting risks. The comparative analysis of model accuracies across these tests is presented in Table 3 in the Results and Discussion section, highlighting the impact of feature selection on model performance.

Model Evaluation Metrics. Model performance was evaluated using multiple metrics, including Accuracy, Precision, Recall, F1-Score, Confusion Matrix, and Receiver-Operating Characteristic Curve (ROC) & Area Under Curve Score (AUC). These evaluations verified the ensemble model's robustness and its predictive accuracy. Also, utilizing multiple evaluation metrics ensures a comprehensive assessment of model performance, essential for critical clinical decision-making scenarios.

Results and Discussions

The synthetic dataset, comprising 1,072 records with 13 features, underwent comprehensive exploratory data analysis (EDA). The histograms and boxplots identified skewness and outliers primarily in features such as Age, Creatinine Phosphokinase, Serum Creatinine, and Platelets. Correlation analysis revealed that the feature Time had a strong negative correlation (-0.55) with the target feature (DEATH_EVENT), emphasizing the importance of follow-up duration in patient outcomes. Identifying such strong correlations aids in model interpretability and clinical decision-making.

Feature selection significantly refined model training by identifying the most impactful features: Age, Ejection Fraction, Serum Creatinine, and Time. These features consistently appeared as the top predictors across Recursive Feature Elimination (RFE), Random Forest, and Information Gain methods, thus justifying their retention despite slight accuracy reductions in certain models during initial testing. This selection aligns well with clinical evidence highlighting these factors as critical predictors of HF survival. The machine learning models were trained across three tests, as shown in Table 3, which presents the accuracies obtained for each model under different conditions.

ML Model	Without feature selection (%)	With feature selection (%)
LR	87	87
RF	93	90
SVM	87	88
XGBoost	93	93
MLP	91	90
Extra Trees	93	92
CatBoost	94	94
LightGBM	94	91

Table 3: Accuracies of ML Models Across Tests

- **Test 1** (without feature selection): Models exhibited high accuracies, with CatBoost and LightGBM achieving the highest (94%), followed closely by Random Forest, XGBoost, and Extra Trees (93%). Simpler models, such Logistic Regression and SVM, displayed lower accuracy (87%) due to their limited ability to capture complex, non-linear relationships.
- **Test 2** (with feature selection): Despite minor accuracy reductions (e.g., Random Forest from 93% to 90%, LightGBM from 94% to 91%), feature selection was retained for interpretability, reduced computational complexity, and mitigation of overfitting. CatBoost maintained consistent performance (94%), underscoring its robustness with fewer features.
- **Test 3** (stacked ensemble model): Based on Test 1 results, CatBoost, LightGBM, and XGBoost were selected as base models, with Multilayer Perceptron (MLP) chosen as the meta-learner due to its strong performance (91%) and capability to capture complex patterns distinct from tree-based models. After hyperparameter optimization using RandomizedSearchCV and cross-validation, the ensemble achieved a high accuracy of 95%.

The final ensemble demonstrated a precision of 98% and recall of 95%, achieving an impressive balance (F1-score of 96%).

The ROC curve, as illustrated in Figure 5, yielded an AUC of 98%, further confirming the model's robustness and strong discrimination ability between survival and non-survival cases. High AUC values emphasize the model's reliability in differentiating patient outcomes across various decision thresholds.

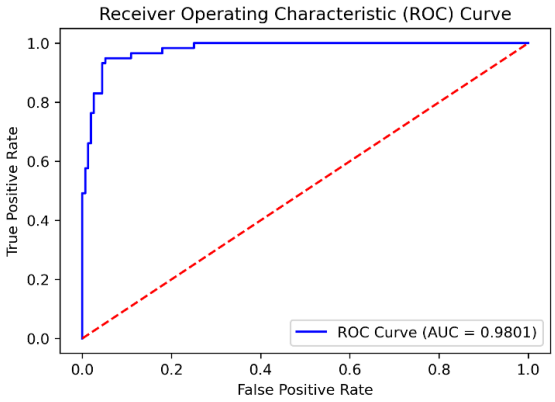

Figure 5: ROC Curve

The confusion matrix, depicted in Figure 6, indicated 148 true positives and 56 true negatives, demonstrating the model's high capability in correctly classifying both deceased and survived patients. Minimal misclassifications were observed, with 8 false positives and 3 false negatives, suggesting a low rate of incorrect survival or mortality predictions. This low misclassification rate is particularly important in medical applications, where false negatives (incorrectly predicting survival for a patient at risk) could lead to inadequate clinical intervention. Furthermore, the high precision (98%) and recall (95%) values reinforce the model's ability to correctly identify patients at risk while minimizing unnecessary false alarms. The model's strong performance metrics suggest that it could serve as an effective decision-support tool

For early intervention, helping healthcare providers prioritize high-risk patients for closer monitoring and timely treatment.

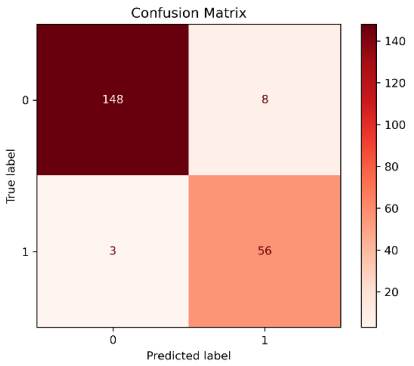

Figure 6: Confusion Matrix

Limitations and Future works. Despite its high accuracy, this study has limitations. The synthetic dataset, though achieved a high similarity (91%) to the original UCI dataset, may still lack fine-grained clinical details, making real-world application challenging. Its focus on binary and continuous features limits the inclusion of complex medical attributes like treatment history or lifestyle factors. Although SMOTE helped address class imbalance, synthetic oversampling can introduce biases that alter the dataset's distribution. The ensemble model achieved 95% accuracy, but the lack of interpretability in deep learning-based meta-learners makes clinical deployment difficult. To enhance this research, future work can explore alternative synthetic data generation methods to improve realism and reduce potential biases introduced by current generation techniques. Incorporating real-world clinical data, such as COVID-19 patient histories and longitudinal health records, could improve model adaptability to evolving medical conditions. Further refinement in feature selection techniques, particularly methods that capture interaction effects and non-linear dependencies, may enhance predictive power while maintaining efficiency. Lastly, improving the interpretability of the ensemble model by integrating Explainable AI (XAI) techniques would ensure that clinical practitioners can better understand and trust the model's predictions, facilitating smoother real-world implementation.

Conclusion

This study investigated heart failure survival prediction using various machine learning models and an ensemble approach. The final stacked ensemble model achieved an accuracy of 95% and an AUC of 98%, significantly outperforming individual models. The use of synthetic data proved effective in overcoming data scarcity, enabling robust predictive modeling. This research highlights the potential of machine learning in healthcare, providing valuable tools for early identification and improved management of high-risk heart failure patients.

References

Ahmad, T., Munir, A., Bhatti, S. H., Aftab, M., and Raza, M. A. "Survival analysis of heart failure patients: A case study." *PLoS ONE*, 12(7), e0181001, Jul. 2017. doi: 10.1371/journal.pone.0181001.

Bataineh, A. A., and Manacek, S. "MLP-PSO hybrid algorithm for heart disease prediction." *Journal of Personalized Medicine*, 12(8), 1208, Jul. 2022. doi: 10.3390/jpm12081208.

Bourazana, A., et al. "Artificial intelligence in heart failure: friend or foe?" *Life*, 14(1), 145, Jan. 2024. doi: 10.3390/life14010145.

Chicco, D., and Jurman, G. "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone." *BMC Medical Informatics and Decision Making*, 20(1), Feb. 2020. doi: 10.1186/s12911-020-1023-5.

Davis, R., Hobbs, F. D. R., and Lip, G. Y. H. "ABC of heart failure: History and epidemiology." *BMJ. British Medical Journal*, 320(7226), 39–42, Jan. 2000. doi: 10.1136/bmj.320.7226.39.

Hajar, R. "Congestive heart failure: A history." *Heart Views*, 20(3), 129, Jan. 2019. doi: 10.4103/heartviews.heartviews_77_19.

Hu, K., Ertl, G., Frantz, S., and Nordbeck, P. "Heart failure classification in clinical practice: time to redefine?" *Chinese Medical Journal*, 135(9), 1039–1040, Jan. 2022. doi: 10.1097/cm9.0000000000001823.

Hunt, S. A., et al. "2009 focused update incorporated into the ACC/AHA 2005 guidelines for the diagnosis and management of heart failure in adults." *Circulation*, 119(14), Apr. 2009. doi: 10.1161/circulationaha.109.192065.

Inamdar, A. A., and Inamdar, A. C. "Heart failure: Diagnosis, management and utilization." *Journal of Clinical Medicine*, 5(7), 62, Jun. 2016. doi: 10.3390/jcm5070062.

Ishaq, A., et al. "Improving the prediction of heart failure patients' survival using SMOTE and effective data mining techniques." *IEEE Access*, 9, 39707–39716, 2021. doi: 10.1109/ACCESS.2021.3064084.

Joloudari, J. H., Marefat, A., Nematollahi, M. A., Oyelere, S. S., and Hussain, S. "Effective class-imbalance learning based on SMOTE and convolutional neural networks." *Applied Sciences*, 13(6), 4006, Mar. 2023. doi: 10.3390/app13064006.

Kapelios, C. J., Shahim, B., Lund, L. H., and Savarese, G. "Epidemiology, Clinical Characteristics and Cause-specific

Outcomes in Heart Failure with Preserved Ejection Fraction." *Cardiac Failure Review*, 9, Nov. 2023. doi: 10.15420/cfr.2023.03.

Kaushik, S., and Birok, R. "Heart failure prediction using Xgboost algorithm and feature selection using feature permutation." *2021 Fourth International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, Sep. 2021. doi: 10.1109/ICECCT52121.2021.9616626.

Lang, R. M., et al. "Recommendations for cardiac chamber quantification by echocardiography in adults: An update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging." *Journal of the American Society of Echocardiography*, 28(1), 1-39.e14, Jan. 2015. doi: 10.1016/j.echo.2014.10.003.

Li, L., et al. "Prediction of hospital mortality in mechanically ventilated patients with congestive heart failure using machine learning approaches." *International Journal of Cardiology*, 358, 59–64, Jul. 2022. doi: 10.1016/j.ijcard.2022.04.063.

Mamun, M., Farjana, A., Mamun, M. A., Ahammed, M. S., and Rahman, M. M. "Heart failure survival prediction using machine learning algorithm: am I safe from heart failure?" *2022 IEEE World AI IoT Congress (AIIoT)*, Seattle, WA, USA, 2022, 194–200. doi: 10.1109/AIIoT54504.2022.9817303.

Masetic, Z., and Subaşi, A. "Congestive heart failure detection using random forest classifier." *Computer Methods and Programs in Biomedicine*, 130, 54–64, Jul.

Moreno-Sánchez, P. A. "Improvement of a prediction model for heart failure survival through explainable artificial intelligence." *Frontiers in Cardiovascular Medicine*, 10, Aug. 2023. doi: 10.3389/fcvm.2023.1219586.

Newaz, A., Ahmed, N., and Haq, F. S. "Survival prediction of heart failure patients using machine learning techniques." *Informatics in Medicine Unlocked*, 26, 100772, Jan. 2021. doi: 10.1016/j.imu.2021.100772.

NHLBI, NIH. "What is heart failure?" *NHLBI, NIH*, Mar. 24, 2022. Available: https://www.nhlbi.nih.gov/health/heart-failure.

NIHR. "Medication used to prevent heart attacks may be ineffective for British South Asians." Available: https://www.nihr.ac.uk/news/medication-used-to-preventheartattacks-may-be-ineffective-for-british-south-asians/34274.

Pocock, S., et al. "Predicting survival in heart failure: A risk score based on 39,372 patients from 30 studies." *European Heart Journal*, 34(19), 1404–1413, Oct. 2012. doi: 10.1093/eurheartj/ehs337.

Ponikowski, P., et al. "2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure." *European Heart Journal*, 37(27), 2129–2200, May 2016. doi: 10.1093/eurheartj/ehw128.

Quinn, C. *100 Questions and Answers about Congestive Heart Failure*. Jones & Bartlett Learning, 2006. Available: https://books.google.ae/books?id=gLKZHdv4KWgC.

Ribatti, D. "William Harvey and the discovery of the circulation of the blood." *Journal of Angiogenesis Research*, 1(1), 3, Jan. 2009. doi: 10.1186/2040-23841-3.

Sabbah, H. N. "Silent disease progression in clinically stable heart failure." *European Journal of Heart Failure*, 19(4), 469–478, Dec. 2016. doi: 10.1002/ejhf.705.

Sang, X., Yao, Q. Z., Ma, L., Cai, H. W., and Luo, P. "Study on survival prediction of patients with heart failure based on support vector machine algorithm." *2020 International Conference on Robots & Intelligent System (ICRIS)*, Sanya, China, 2020, 636–639. doi: 10.1109/ICRIS52159.2020.00160.

Saravanan, S., and Swaminathan, K. "Hybrid K-Means and Support Vector Machine to predict heart failure." *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*, Trichy, India, 2021, 1678–1683. doi: 10.1109/ICOSEC51865.2021.9591738.

Sarra, R. R., Dinar, A. M., and Mohammed, M. A. "Enhanced accuracy for heart disease prediction using artificial neural network." *Indonesian Journal of Electrical Engineering and Computer Science*, 29(1), 375, Jan. 2022. doi: 10.11591/ijeecs.v29.i1.pp375-383.

Savarese, G., Becher, P. M., Lund, L. H., Seferovic, P., Rosano, G., and Coats, A. J. S. "Global burden of heart failure: a comprehensive and updated review of epidemiology." *Cardiovascular Research*, 118(17), 3272–3287, Feb. 2022. doi: 10.1093/cvr/cvac013.

Sciomer, S., et al. "Role of gender, age and BMI in prognosis of heart failure." *European Journal of Preventive Cardiology*, 27(2_suppl), 46–51, Dec. 2020. doi: 10.1177/2047487320961980.

Siriwardena, M., and Fan, E. "Descent into heart and lung failure." In *Elsevier eBooks*, 2018, pp. 3–36. doi: 10.1016/b978-0-12-810491-0.00001-1.

Tufail, S., Riggs, H., Tariq, M., and Sarwat, A. I. "Advancements and challenges in machine learning: A comprehensive review of models, libraries, applications, and algorithms." *Electronics*, 12(8), 1789, Apr. 2023. doi: 10.3390/electronics12081789.

Virani, S. S., et al. "Heart Disease and Stroke Statistics—2021 update." *Circulation*, 143(8), Feb. 2021. doi: 10.1161/cir.0000000000000950.

Zheng, C., et al. "Hierarchical management of chronic heart failure: a perspective based on the latent structure of comorbidities." *ESC Heart Failure*, 9(1), 595–605, Nov. 2021. doi: 10.1002/ehf2.13708.

Zuin, M., Rigatelli, G., Roncon, L., Pasquetto, G., and Bilato, C. "Risk of incident heart failure after COVID-19 recovery: a systematic review and meta-analysis." *Heart Failure Reviews*, 28(4), 859–864, Dec. 2022. doi: 10.1007/s1074102210292-0.