Cardiac Failure Forecasting

Subjects: Computer Science, Artificial Intelligence

Contributor: Istiak Mahmud , Md Mohsin Kabir , M. F. Mridha , Sultan Alfarhood , Mejdl Safran , Dunren Che

Accurate prediction of heart failure can help prevent life-threatening situations. Several factors contribute to the risk of heart failure, including underlying heart diseases such as coronary artery disease or heart attack, diabetes, hypertension, obesity, certain medications, and lifestyle habits such as smoking and excessive alcohol intake. Machine learning approaches to predict and detect heart disease hold significant potential for clinical utility but face several challenges in their development and implementation.

cardiac failure metamodel forecasting

1. Introduction

Heart failure is a complex and potentially life-threatening condition that significantly burdens healthcare systems worldwide. It is a pathophysiologic condition in which the heart's inability to pump blood at a rate sufficient to meet the needs of the body's metabolizing tissues results from faulty cardiac function [1][2]. It includes a number of heart-related illnesses, such as coronary artery disease, heart attacks, heart failure, arrhythmias, and several other cardiovascular ailments. Heart disease is a leading cause of death globally [3], accounting for many premature deaths and posing a significant burden on healthcare systems. Heart disease is a common and significant health issue in many parts of the world [4]. The American Heart Association says that heart failure is projected to increase dramatically [5]. Accurate prediction of heart failure can play a vital role in early detection and prevention of adverse outcomes, ultimately leading to improved patient outcomes and reduced healthcare costs.

Timely and accurate detection of heart failure is crucial for effective management and treatment [6]. Detecting heart failure early allows for prompt intervention and the implementation of appropriate medical strategies, which can help slow the progression of the disease, alleviate symptoms, and improve the patient's quality of life. Early detection can also reduce the risk of complications and hospitalizations associated with advanced stages of heart failure. From 1989 [7] to now, there have been many approaches to finding the best methods for cardiac failure prediction. In 2017, Simge et al. [8] used Matlab and WEKA to find the best way to detect heart failure disease and obtained a good accuracy of 67.7% for the ensemble subspace discriminant algorithm and the decision tree algorithm. Then, in 2018, Ali et al. [9] utilized the Claveland dataset [10] for their studies and obtained 84% accuracy for the Naive Bayes algorithm. Further, in 2019, Saba et al. [11] performed prediction for heart diseases and obtained 84.85% accuracy for the logistic regression (SVM) technique. However, most of them used the same dataset from the UCI repository [12], which contains 300 records. This is a rather limited amount of data for machine learning training.

Machine learning techniques have drawn a lot of attention in the medical field lately because of their potential to help with the detection and prediction of cardiac disease [13]. Large volumes of clinical data may be analysed by machine learning algorithms to find links and patterns that are not immediately obvious to human practitioners [14]. These algorithms can harness the power of computer models to make accurate predictions and provide valuable insights into disease risk assessment. However, the development and implementation of machine learning models for heart failure prediction face several challenges [15]. The complexity of the cardiovascular system and the multifactorial nature of heart failure necessitate integrating diverse data sources, including clinical test data, medical imaging, and patient demographics. Data quality, feature selection, and model performance issues must be addressed to ensure reliable and clinically relevant predictions.

2. Cardiac Failure Forecasting

Heart failure forecasting has garnered significant attention recently due to its potential to enhance patient care and improve healthcare resource allocation. Numerous studies have explored the application of machine learning and deep learning techniques in predicting the onset and progression of heart failure. These methods leverage the abundance of clinical and physiological data available, aiming to provide early and accurate prognostic insights for clinicians and patients. Researchers review the literature on heart failure forecasting, focusing on the various machine learning and deep learning approaches employed, the datasets utilized, and the reported performance metrics.

Liang et al. [16] proposed a novel deep learning model called tBNA-PR to accurately predict heart failure and identify sub-phenotypes using temporal electronic health records (tEHRs) data. The model effectively captures the complexity and heterogeneity of the data to obtain informative patient representations. The study demonstrates the effectiveness of tBNA-PR on a real-world dataset, achieving prediction accuracy of 0.78, F1-Score of 0.7671, and AUC of 0.7198, outperforming existing benchmarks. The analysis identifies three distinct sub-phenotypes of heart failure patients based on clustering and subgroup analysis, revealing specific characteristics and significant features associated with each sub-phenotype. The findings have practical implications for clinical decision support, but the study acknowledges limitations related to data completeness, disease specificity, generalizability, interpretability, and the need for further research.

In a study by Robert et al. [7], a novel algorithm was proposed for diagnosing coronary artery disease, employing a probability-based approach. This algorithm's reliability and clinical utility were tested across three patient test groups. 303 consecutive patients who were sent for coronary angiography at the Cleveland Clinic between May 1981 and September 1984 served as the reference group for the model's development. The study's findings showed that when applied to individuals with chest pain syndromes and intermediate disease prevalence, discriminant functions used to determine coronary disease probabilities produced accurate and clinically helpful results. In another study by Simge et al. [8], a comparison was made between two prominent machine learning platforms using the same dataset. The researchers conducted experiments to classify heart disease using six distinct algorithms: Quadratic SVM, Linear SVM, Cubic SVM, Decision Tree, Medium Gaussian SVM, and Ensemble Subspace Discriminant. These experiments were carried out in both the Matlab environment and WEKA.

The dataset utilized in this study was acquired from the machine learning repository of UCI [12]. The highest accuracy achieved was 67.7% using the Ensemble Subspace Discriminant algorithm in Matlab, while the Decision Tree algorithm in the WEKA platform also yielded an accuracy of 67.7%.

Li et al. [17] introduce a deep learning-based automatic system for diagnosing heart failure by tackling the issue of imbalanced data in chest X-ray (CXR) images. The approach combines under-sampling and instance selection techniques to maintain the integrity of data distribution and presents a comprehensive multi-level classification method to diagnose specific heart failure causes. Experimental results demonstrate that the proposed approach outperforms traditional under-sampling methods, achieving an accuracy of 84.44% in multi-class classification tasks. Rao et al. [18] presented a deep-learning framework for predicting heart failure incidence using electronic health records. The authors developed a novel Transformer-based risk model incorporating patient diagnoses, medications, age, and calendar year. The model achieved high predictive performance, outperforming existing deep learning models. Ablation analysis revealed the importance of medications and calendar year in predicting HF risk. Contribution analyses identified both established risk factors and new associations, providing insights for data-driven risk factor identification. The study highlights the potential of the deep learning model to inform preventive care and identify new hypotheses for further research and drug repurposing studies in HF prediction and other complex conditions.

In a related study conducted by Ali et al. [9], the Cleveland dataset was employed for analysis, and a feature selection process was carried out to train three distinct classifiers, namely Support Vector Machine (SVM), Naïve Bayes, and K-Nearest Neighbors, utilizing a 10-fold cross-validation technique. Their findings revealed that the Naïve Bayes classifier exhibited superior performance on this dataset and the selected features, surpassing or equaling the performance of SVM and KNN across all four evaluation parameters. Notably, it achieved an accuracy rate of 84%. Saba et al. conducted a study [11] that explores the prediction of heart disease using data science methodologies. Their research focuses on employing feature selection techniques and algorithms to improve the accuracy of heart disease prediction. Multiple heart disease datasets were utilized for experimentation and analysis purposes. The authors employed various feature selection techniques, including Logistic Regression, Decision Tree, Random Forest, Nave Bayes, and Logistic Regression SVM, using Rapid Miner as the tool. Notably, the highest achieved accuracy of 84.85% was obtained using the UCI dataset [12] in combination with the Logistic Regression (SVM) technique.

Earlier research on predicting heart failure has mostly relied on two widely recognized datasets, namely the UCI repository and the Cleveland dataset, as indicated in **Table 1**. However, these datasets suffer from limitations in terms of the number of records available for machine learning training purposes. Additionally, prior investigations primarily employed basic machine learning models for detection or forecasting tasks. To address these limitations, researchers conducted the research using a comprehensive dataset comprising 918 records and introduced a novel metamodel for predicting cardiac failure in patients. The metamodel represents a fusion of four distinct machine-learning models, allowing for enhanced accuracy and robustness in forecasting outcomes.

Table 1. A Comparative Overview of Heart Disease Prediction Methods: Various methods and their performance on different datasets, including limitations.

Ref	Method	Dataset	Records Count	Outcome	Limitations
[19]	SVM, NN, Fuzzy Genetic, CART, and Random Forest	Custom dataset	136	CART was most effective in determining the type and degree of heart failure.	Small sample size prevented the proposed model from generalizing well. In determining severity, accuracy is fairly poor.
[20]	Modified Self Adaptive Bayesian algorithm (MSABA) and IoT	UCI	303	Used the pulse sensor of a smart watch to obtain data and find the disease.	Used only ECG data and blood pressure data for analysis and used fewer records for training.
[21]	Five active learning multi-label selection methods: random, MMC, adaptive, AUDI, and Quire	UCI	303	Accuracy in the generalisation of the learning model beyond the available data for the optimised label ranking model.	Fewer data were utilized for prediction and fine-tuning because the system is sophisticated.
[22]	Support Vector Machine (SVM)	UCI and Cleveland datasets	573	The performance of the proposed model was subsequently validated by comparing it to conventional models in 2022 using a number of performance criteria, and the componential load was cut in half.	The features used in the system were decreased from 14 to 6 to reduce the computational load, but sometimes the reduced features are also important for disease analysis.
[23]	Sine Cosine Weighted K-Nearest Neighbour (SCA_WKNN) algorithm	UCI	303	In comparison to WK-NN and K-NN, CA_WKNN obtains maximum accuracy of 4.59% and 15.61%, respectively. Blockchain-powered decentralised storage exceeds peer-to-peer storage in terms of maximum throughput by 25.03 percent.	The operational cost is contingent upon the number of transactions conducted within the system, making it expensive. As the data expands, it becomes necessary to limit the system’s learning capacity to a restricted data source to avoid incurring additional costs.
[24]	LR, ANN, SVM(support	Korea National Health and	6170	Out of the three models applied, the SVM gives	As for the data collection using smartwatch

Ref	Method	Dataset	Records Count	Outcome	Limitations
	vector machine)	Nutrition Examination Survey by smartwatches		the highest accuracy of 83.04% for six types of heart disease prediction.	sensors, all the necessary feature data are not available for proper heart disease prediction.

1. Denolin, H.; Kuhn, H.; Krayenbuehl, H.; Loogen, F.; Reale, A. The defintion of heart failure. Eur. Heart J. 1983, 4, 445–448.
2. Gheorghiade, M.; De Luca, L.; Fonarow, G.C.; Filippatos, G.; Metra, M.; Francis, G.S. Pathophysiologic targets in the early phase of acute heart failure syndromes. Am. J. Cardiol. 2005, 96, 11–17.
3. Ponikowski, P.; Anker, S.D.; AlHabib, K.F.; Cowie, M.R.; Force, T.L.; Hu, S.; Jaarsma, T.; Krum, H.; Rastogi, V.; Rohde, L.E.; et al. Heart failure: Preventing disease and death worldwide. ESC Heart Fail. 2014, 1, 4–25.
4. von Haehling, S.; Anker, S.D. Prevalence, incidence and clinical impact of cachexia: Facts and numbers—Update 2014. J. Cachexia Sarcopenia Muscle 2014, 5, 261–263.
5. Heart, A. Heart failure projected to increase dramatically, according to new statistics. Am. Heart Assoc. News 2017, 1, 2018–2020.
6. Goff, D.C., Jr.; Brass, L.; Braun, L.T.; Croft, J.B.; Flesch, J.D.; Fowkes, F.G.; Hong, Y.; Howard, V.; Huston, S.; Jencks, S.F.; et al. Essential features of a surveillance system to support the prevention and management of heart disease and stroke: A scientific statement from the American Heart Association Councils on Epidemiology and Prevention, Stroke, and Cardiovascular Nursing and the Interdisciplinary Working Groups on Quality of Care and Outcomes Research and Atherosclerotic Peripheral Vascular Disease. Circulation 2007, 115, 127–155.
7. Detrano, R.; Janosi, A.; Steinbrunn, W.; Pfisterer, M.; Schmid, J.J.; Sandhu, S.; Guppy, K.H.; Lee, S.; Froelicher, V. International application of a new probability algorithm for the diagnosis of coronary artery disease. Am. J. Cardiol. 1989, 64, 304–310.
8. Ekız, S.; Erdoğmuş, P. Comparative study of heart disease classification. In Proceedings of the 2017 Electric Electronics, Computer Science, Biomedical Engineerings' Meeting (EBBT), Istanbul, Turkey, 20–21 April 2017; IEEE: New York, NY, USA, 2017; pp. 1–4.
9. Nassif, A.B.; Mahdi, O.; Nasir, Q.; Talib, M.A.; Azzeh, M. Machine learning classifications of coronary artery disease. In Proceedings of the 2018 International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI-NLP), Pattaya, Thailand, 15–18 November 2018; IEEE: New York, NY, USA, 2018; pp. 1–6.

10. UCI Machine Learning Repository: Statlog (Heart) Data Set. Available online: https://archive.ics.uci.edu/ml/datasets/statlog+(heart) (accessed on 1 June 2023).
11. Bashir, S.; Khan, Z.S.; Khan, F.H.; Anjum, A.; Bashir, K. Improving heart disease prediction using feature selection approaches. In Proceedings of the 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 8–12 January 2019; IEEE: New York, NY, USA, 2019; pp. 619–623.
12. UCI Machine Learning Repository: Heart Disease Data Set. Available online: https://archive.ics.uci.edu/ml/datasets/heart+disease (accessed on 1 June 2023).
13. Learning, M. Heart disease diagnosis and prediction using machine learning and data mining techniques: A review. Adv. Comput. Sci. Technol. 2017, 10, 2137–2159.
14. Awan, S.E.; Sohel, F.; Sanfilippo, F.M.; Bennamoun, M.; Dwivedi, G. Machine learning in heart failure: Ready for prime time. Curr. Opin. Cardiol. 2018, 33, 190–195.
15. Ahsan, M.M.; Siddique, Z. Machine learning-based heart disease diagnosis: A systematic literature review. Artif. Intell. Med. 2022, 128, 102289.
16. Liang, Y.; Guo, C. Heart failure disease prediction and stratification with temporal electronic health records data using patient representation. Biocybern. Biomed. Eng. 2023, 43, 124–141.
17. Li, D.; Zheng, C.; Zhao, J.; Liu, Y. Diagnosis of heart failure from imbalance datasets using multi-level classification. Biomed. Signal Process. Control 2023, 81, 104538.
18. Rao, S.; Li, Y.; Ramakrishnan, R.; Hassaine, A.; Canoy, D.; Cleland, J.; Lukasiewicz, T.; Salimi-Khorshidi, G.; Rahimi, K. An explainable Transformer-based deep learning model for the prediction of incident heart failure. IEEE J. Biomed. Health Inform. 2022, 26, 3362–3372.
19. Guidi, G.; Pettenati, M.C.; Melillo, P.; Iadanza, E. A machine learning system to improve heart failure patient assistance. IEEE J. Biomed. Health Inform. 2014, 18, 1750–1756.
20. Subahi, A.F.; Khalaf, O.I.; Alotaibi, Y.; Natarajan, R.; Mahadev, N.; Ramesh, T. Modified Self-Adaptive Bayesian Algorithm for Smart Heart Disease Prediction in IoT System. Sustainability 2022, 14, 14208.
21. El-Hasnony, I.M.; Elzeki, O.M.; Alshehri, A.; Salem, H. Multi-label active learning-based machine learning model for heart disease prediction. Sensors 2022, 22, 1184.
22. Sarra, R.R.; Dinar, A.M.; Mohammed, M.A.; Abdulkareem, K.H. Enhanced heart disease prediction based on machine learning and χ2 statistical optimal feature selection model. Designs 2022, 6, 87.
23. Hasanova, H.; Tufail, M.; Baek, U.J.; Park, J.T.; Kim, M.S. A novel blockchain-enabled heart disease prediction mechanism using machine learning. Comput. Electr. Eng. 2022, 101, 108086.

24. Kim, M.J. Building a cardiovascular disease prediction model for smartwatch users using machine learning: Based on the Korea national health and nutrition examination survey. Biosensors 2021, 11, 228.

Retrieved from https://encyclopedia.pub/entry/history/show/108647