JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

HARNESSING ADVANCED DEEP LEARNING TECHNIQUES FOR CARDIAC ILLNESS PREDICTION

[1]Asad Choudhary, [2]Vishnu Potdar, [2]Abhishek Bhatt, [2] Shubhangi Tidake
[1]PG Scholar, [2]Faculty
[1]School of Data Science,
[1]Symbiosis Skills and Professional University, Pune, India

Abstract : Cardiac illness is a major cause of death worldwide, and lifestyle factors such as diet, physical activity and drug use play a crucial role in the progression of the disease. Initial diagnosis and management are important to improve patient outcomes and reduce treatment burden. Medical records of 299 patients with heart problems were examined. These data, which include demographic information, medical history, and medical assessments, are intended to develop a predictive model to classify patients with a medical history of left ventricular dysfunction and heart failure. Using the minority variance method (SMOTE). Evaluate model effectiveness using metrics like accuracy, precision, recall, F1 score and ROC AUC score. The results prove that the 1D-CNN model outperforms deep learning models when used with autoencoders, demonstrating the effectiveness of this approach in predicting cardiac survival. The ability of learning models, especially when augmented with autoencoders, can improve the prediction and management of heart failure. This can lead to significant improvements in patient outcomes and healthcare decisions.

***IndexTerms* - autoencoders, cardiac illness, cardiovascular diseases, Convolutional Neural Network (CNN), deep learning, healthcare, neural networks, predictive model, SMOTE.**

I. INTRODUCTION

Cardiac illness, also known as heart disease, is a condition characterized by the heart's inability to effectively pump blood, resulting in inadequate delivery of oxygen and nutrients to the body's cells. It is a form of heart disease and the major cause of death worldwide. Heart Disease is responsible for approximately 32% of global deaths each year, with most occurring prematurely, before age 70. Diet, lack of exercise, smoking and alcohol consumption. These conditions can contribute to an increased risk of cardiac illness, including high blood pressure, high blood sugar, and obesity. It is often caused by an underlying condition that affects heart function, highlighting the importance of early detection, identification, and supervision. Autoencoders for deep learning, Standard neural network, and one-dimensional convolutional neural networks (1D-CNN) were used to analyze medical data of patients with cardiac problems. The data used in this study consisted of data from 299 patients, including demographic characteristics, medical history, and medical evaluations [1]. Population uses the oversampling technique (SMOTE) to balance the dataset [2]. This process creates a synthetic model from a limited class, improving the model's adeptness at accurately identifying patients based on their medical data. History of systolic dysfunction and heart failure. The model's performance will be evaluated by measuring accuracy, accuracy, recall, and F1 score, with the aim of improving early diagnosis and cardiac management failure, improving patient outcomes, and reducing the treatment burden of cardiovascular disease.

II. LITERATURE REVIEW

Cardiac illness is a progressive disease triggered by the heart's incompetence to circulate blood adequately, resulting in a decreased heart rate. It is classified as a cardiovascular disease (CVD), a major contributor to global mortality, responsible for an estimated 17.9 million deaths annually across both developed and developing nations. [3], [4]. This condition is usually caused by changes in the muscles in the heart (especially the left ventricle) that cause the heart to not pump fully [5]. A significant role that focuses on the timing of a specific event, such as a death. Ahmad and colleagues [1] used Kaplan-Meier and Cox regression to investigate overall survival in patients with class III/IV Cardiac failure. Newaz [6] utilized a Random Forest classifier to predict survival in the same dataset, achieving a maximum G-mean score of 76.83% and a sensitivity score of 80.21%. Bayesian methods have also been adopted, and Ashine and colleagues [7] used Bayesian methods to assess the probability of medical events with better results than other methods. Factors such as ejection fraction (EF), serum creatinine level, age, hypertension, and anemia. Rahul Katarya [8] described the use of certain supervised machine learning algorithms for predicting heart problems. Riddi Kasabe [9] have evaluated various classification techniques for diagnosing heart cardiac illness in patients. Zheng [10] used the model named LightGBM to predict all-cause mortality in chronic Cardiac patients. Chicca [11] carried out an aggregate-analysis and established a survival predictor based on EF and serum creatinine levels. Decision tree, SVM, random forest, XGBoost and LightGBM algorithms give satisfactory results. Moreno [12] compared various models and found that additional classification trees were the most efficient and best interpreted. Sabastian and colleagues [13] used ML to diagnose Cardiac illness with preserved ejection fraction, with AUC of 0.89. Amit Tak [14] employed machine learning classifiers to predict cardiac illness survival in 299 patients, achieving an accuracy of 89.5%, AU-ROC of 93%, sensitivity of 87%, and specificity of 92%, with predictors including hypertension, age, creatinine concentration, CPK, ejection fraction, and sodium concentration, showcasing improved performance metrics. Noor Basha [15] utilized various machine learning algorithms to predict and analyze the presence of cardiac illness in patients at an early stage of illness. They compared these machine learning algorithms with recent literature and found that the K-Nearest Neighbors (KNN) algorithm is the most efficient model. Vosough and colleagues [16] compared six algorithms for predicting cardiac illness readmissions and used random forest as the best performer. Kwon [17] implemented deep learning and ML systems to predict the mortality of patients, with the deep model performing the best. Chu-Hsing Lin [18] compared Convolutional Neural Networks (CNN) and regular Neural Networks (NN) for predicting heart disease using the Cleveland dataset and their findings revealed that the CNN model was more stable. Tulay Karayilan [19] emphasize the importance of early cardiac illness diagnosis and have proposed a model trained using an artificial neural network (ANN). These models need further investigation and validation to ensure they are valid and applicable in medical settings.

III. DATA

The research study uses the data obtained by Faisalabad Heart Institute which included 299 cardiac patients' Medical History [1]. Between the ages of 40 and 95, with a history of diagnosed heart disease, left ventricular systolic dysfunction, and heart failure, diabetes, ejection fraction, gender, platelets, serum creatinine, serum sodium, smoking, duration, and mortality. Five of the characteristics were binary: anemia, hypertension, diabetes, gender, and smoking. The patient did not survive. This inconsistency in the training process can lead to bias in the model. Synthetic Minority Oversampling Technique (SMOTE) is utilized to counter the imbalance issue. SMOTE is a profile augmentation technique that creates a composite data set for small classes (in this case, non-live events). After using SMOTE, the dataset was balanced with 171 files in both survival and non-survival categories [20].

IV. METHODOLOGY

In this section, an overview of the steps involves in deep learning techniques explained, including data loading, analysis, feature selection, data segmentation, model design and training, benchmarking, handling class variances, and standardized testing.

4.1 Data Loading

The medical dataset for patients with heart failure is stored in a pandas DataFrame. This information is necessary to understand the factors that influence heart failure and develop predictive models to improve patient outcomes. Loading data into Pandas DataFrame makes it easy to explore the properties and distribution of datasets, identify missing or erroneous values, and prioritize data for the next review. This step is crucial to guarantee the quality and integrity of the data prior to algorithm utilization.

4.2 Data Analysis

This dataset has 13 variables and 299 lines; there are no missing cells or duplicate rows. It occupies 30.5 KB of memory and has an average line size of 104.4 bytes. On the contrary, 7 is numerical and 6 is categorical. Table I provides a brief summary of the size, completeness, and composition of the data set, forming the basis for further analysis and modeling, "Smoking", "Anemia", "Diabetes", "Hypertension" and "Death". Each variable has two unique values, represented as 1 or 0, representing a binary group. For example, in the "gender" variable, 1 may represent male and 0 may represent female. The value count column shows the frequency of each value in the data. To visualize this data, a bar chart can be drawn for each variable, where the x-axis represents a specific value (1 or 0), and the y-axis represents the frequency.

TABLE I. statistical summary for qualitative variables

Variables	**Metrics**	
	No. of Unique Values	***Value Counts***
sex	2	{1: 194, 0: 105}
smoking	2	{0: 203, 1: 96}
anaemia	2	{0: 170, 1: 129}
diabetes	2	{0: 174, 1: 125}
high_blood_pressure	2	{0: 194, 1: 105}
DEATH_EVENT	2	{0: 203, 1: 96}
sex	2	{1: 194, 0: 105}

“Fig. 1”, below provides a graphical representation of the distribution of variables for each group, showing the relative importance of each group.

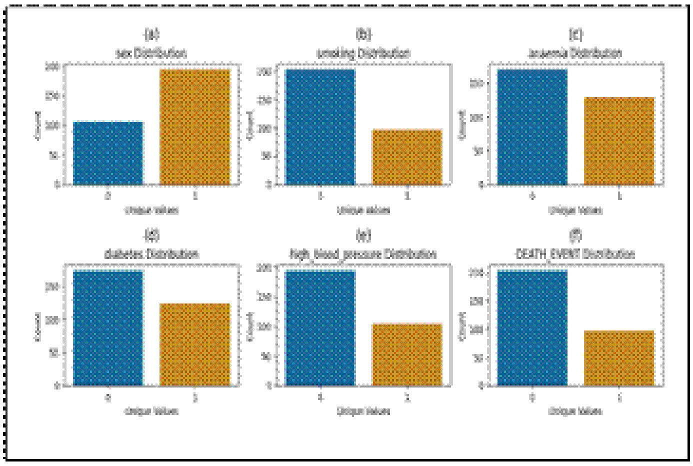

Fig. 1. graphical representation of qualitative variables distribution

Below Table II. Gives statistical insights on the dataset. The age column shows that the arithmetic mean age lies around 60.83 years and its dispersion around 11.89. This indicates a moderate distribution of ages around the athematic mean. The “Creatinine_Phosphokinase " column has an arithmetic mean around 581.84 and its dispersion of 970.29, thus indicating fluctuations between patients. The “Platelets” column shows a mean of approximately 263.358 and a standard deviation of 97.804, indicating the variability of platelet counts among patients. Skewness and kurtosis values give an idea about the shape of the distribution; Skewness measures symmetry and kurtosis measures tail. By dividing the dispersion by the mean, coefficient of variation is calculated which expresses difference between variables as a percentage. These statistics collectively offer an overview of the data set's numerical values, aiding in understanding its distribution and evolution.

TABLE II. statistical summary for quantitative variables

Variables	**Metrics**			
	Mean	***Standard Deviation***	***Sum***	***Variance***
age	60.83	11.89	18189.33	141.48
creatinine_phosphokinase	581.83	970.28	173970	941458.5
ejection_fraction	38.08	11.83	11387	140.06
platelets	263358	97804.2	78744050	9565668749
serum_creatinine	1.39	1.03	416.77	1.07
serum_sodium	136.62	4.41	40851.0	19.46
time	130.26	77.61	38948.0	6023.96

Corresponding to these statistics, below “Fig. 2”, show the shape and distribution of the data, providing a representation of the distribution of results for each variable.

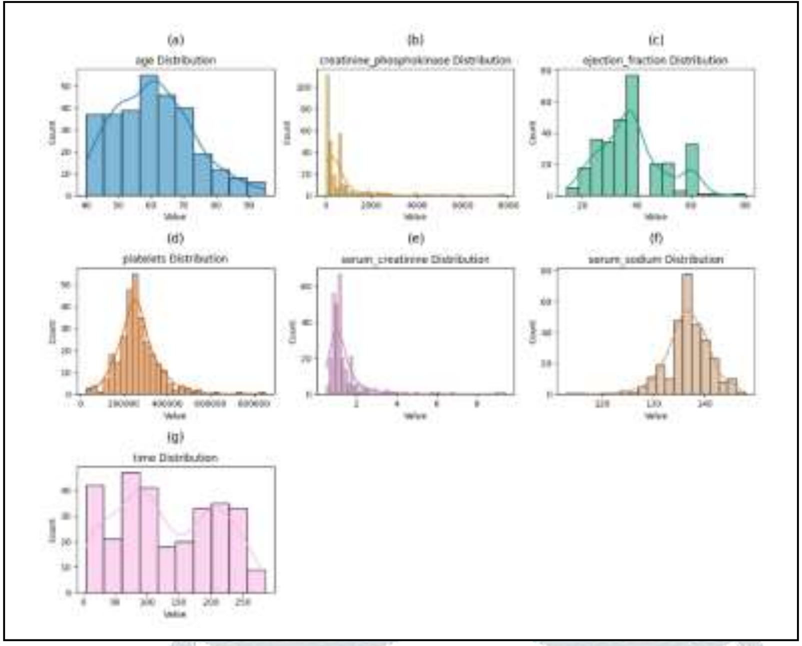

Fig. 2. graphical representation of quantitative variables distribution

4.3 Feature Selection

The most crucial factors in predicting cardiac illness, are time, age, creatine phosphokinase, and blood creatinine, which are important in predicting cardiac illness. These features contain valuable information that may affect the results, such as how long the patient was monitored, his age, and the level of some drugs in his blood. Smoking, high blood pressure, diabetes, and blood sodium. These features can provide additional detail to help enhance the comprehension of a patient's health by incorporating a broader range of factors and metrics related to heart problems. Therefore, three features were chosen to predict the DEATH_EVENT target variable: time, serum creatinine, and age.

4.4 Data Splitting

To prepare the data for model training and evaluation, we split the data into three groups using the train_test_split function in sklearn.model_selection: training (70%), testing (15%), and validation (15%). This approach ensures that the model is trained on ample amounts of data samples, tested on unseen data samples to evaluate its performance, and validated on different data to improve its performance and avoid overwork. This data segmentation strategy helps make the model more general and useful in predicting survival.

4.5 Handling Class Imbalance

We use the mixed minority sampling method implemented by the "SMOTE" module to counter the issue of class bias in the training process [2]. Class imbalance occurs when one class (for example, patients with heart disease) is underrepresented compared to another class (for example, patients without the disease), which make the sample can look bad. SMOTE works by creating a synthetic model with a goal of maintaining the small number class distribution, thus increasing the model's ability to grasp from a small number of classes and making more accurate predictions. The training process ensures that the model does not show bias for most groups and can learn well from all groups. This approach can enhance the model's performance, particularly when the class distribution is heavily skewed, which could otherwise impact the model's capacity to extend towards new data.

4.6 Feature Scaling

We used the "sklearn.preprocessing.StandardScaler" standard to ensure that all features have the same value and the same scale in the model. Standardizing the data to have a mean of 0 and a standard deviation of 1 for each feature. It also improves the interpretation of model coefficients because the scale of the features is now the same. This initial step ensures that the model can effectively learn from the data and generate precise predictions.

4.7 Model Building and Training

In this study, we use a neural network, autoencoder, and 1D CNN algorithms to construct and train our predictive model. These algorithms were selected for their performance in processing sequence data and providing relevant features for predicting the outcome of heart failure.

4.7.1 Neural Network

A neural network (NN) is a computational framework inspired by the structure and function of the human brain. It comprises nodes or neurons organized in a network that processes input data to generate output. Each neuron applies weights to the input, processes it through an activation function, and passes the result to the next layer. In Equation (1), f is the activation function, w_i are the weights, x_i are the inputs, and b is the bias term. The function allows it to learn complex patterns in data by detecting inequalities in patterns. In a neural network, multiple layers of neurons are stacked on top of each other; Each layer in the neural network accepts input from the previous layer and transfers its output to the subsequent layer for processing, resulting in a final prediction or classification. In the training process, the network adapts its weights and biases to improve performance, typically using optimization methods like stochastic gradient descent (SGD) or Adam. In this specific model, there are three layers: two "Dense" layers with 64 units each and ReLU activation functions, and an output layer with 1 unit and a sigmoid activation function for binary classification.

The binary cross-entropy loss function is used for optimization. To prevent overfitting, early stopping and model checkpoint callbacks are utilized, tracking validation loss, and saving the best model weights. The model is trained on batches of data and validated on a separate validation set. After training, the model is evaluated on the test set by estimating class probabilities and transforming them into binary predictions using a threshold of 0.5. Finally, a report is generated to assess the model's performance on the test set.

$$y = f \sum_{i=1}^{n} w_i \cdot x_i + b \qquad (1)$$

4.7.2 1D Convolutional Neural Network

In the realm of neural networks, a one-dimensional convolutional neural network (CNN) model is especially effective in handling continuous data such as time series or sequential data. Unlike traditional neural networks that process entire input data, 1D CNNs utilize layers to capture local features from the network. The Slide filters or kernels over the input data to make connections and capture patterns such as edges, shapes, or time dependency. The results of the convolutions are then put through a smoothing and blending process to reduce the size and preserve key features. This layering technique allows 1D CNNs to learn relevant features from data sequences, making it ideal for tasks such as speech recognition, emotion recognition, and physical activity. Frequently enhance the loss using optimization techniques such as stochastic gradient descent (SGD) or Adam. And use ReLU as a function. This layer is followed by a Flattening layer to convert the 3D output to 1D and two dense layers of 64 and 1 units using ReLU and sigmoid activation functions for the hidden and output layers, respectively. Throughout training, the model is tuned using the Adam optimizer and the binary cross-entropy loss function. Additionally, early stop and sample check callbacks are defined to check validation error and record sample weighting accordingly. The model is trained on the training set with a size of 32 sets and there is confirmation of the validation process. The model is then assessed on the test set by estimating the class probability and transforming it into a binary predictor using a threshold value of 0.5. Finally, a distribution map is created to evaluate the standard performance of the test system. Perform early stopping and model checkpointing to improve training efficiency and save the best model.

4.7.3 Autoencoder

Autoencoders are techniques used for unsupervised learning and dimensionality reduction. It works by training a neural network to understand details of input data, called the latent space, and then reconstructing the original input based on that representation. Autoencoders include encoders and decoders. The encoder reduces input data into the hidden space, and the decoder rebuilds the original input from this condensed representation. By minimizing reconstruction errors, autoencoders can learn useful information representations that capture key features. This method is often applied in activities such as image noise reduction, outlier detection, and attribute extraction, where it is beneficial for grasping a summarized view of the input data. The layers form the input into two hidden parts, and the decoder process reconstructs the original input. Autoencoders are trained to minimize the mean square error of the input and reproduce the output with early stop and callback test patterns. The encoded features serve as input to a neural network classifier, which typically includes a first layer, a hidden layer, and final layer. The Neural network algorithm was trained to find encoded features using the Binary cross-entropy loss function and Adam optimizer technique, and similar calls were used to improve the training. Finally, the training model is evaluated by predicting the outcome of the test and creating an evaluation to evaluate its performance.

4.8 Model Evaluation

In model evaluation, several key metrics are often used to evaluate the effectiveness of classification models:

4.8.1 Precision

Precision measures the proportion of correctly found features of each event that are predicted correctly. It evaluates the model's adeptness at avoid defects and can be formulated as shown in Equation (2).

$$Precision = \frac{True\,Positives}{True\,Positives + False\,Positives} \qquad (2)$$

4.8.2 Recall

Recall is a proportional measure that finds the proportion of correct positive events among all positive events. It evaluates the model's adeptness at capture all positive factors and given as in Equation (3).

$$Recall = \frac{True\,Positives}{True\,Positives + False\,Negatives} \qquad (3)$$

4.8.3 F1 Score

The F1 score is a compromise between precision and recall and provides a measure of the balance between the two parameters. It calculated as shown in Equation (4).

$$\text{F1 Score} = 2 * \frac{\Pr ecision * Recall}{\Pr ecision + Recall} \qquad (4)$$

4.8.4 Matthew’s Correlation Coefficient

A balanced measure of sample performance by accounting for false positives and true positives. It calculated as shown in Equation (5).

$$\frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)}} \qquad (5)$$

4.8.5 ROC AUC Score

The ROC AUC score measures the area under the ROC curve and shows the model's adeptness at distinguishing between positive and negative variables.

4.8.6 Average Precision

Average precision is the area under the true recall curve and provides an added measure for assessing the model's ability to represent variance accurately and precisely.

4.8.7 Accuracy

The correct proportion of identified events out of the total events is known as accuracy. It is formulated as shown in Equation (6).

$$Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Events} \qquad (6)$$

V. RESULTS AND ANALYSIS

Table III, below shows performance metrics for three models: Neural Network, 1D CNN, and Autoencoder. The 1D CNN performs best in most metrics, focus, recall, ROC AUC score and F1 score. This shows that the 1D CNN can identify good examples and capture most of the actual good examples, making it a viable choice. Although the deep learning autoencoder model still performs well, it lags behind the two models regarding precision, recall, ROC AUC score and F1 score. This concludes, although autoencoder may perform well on the data, its prediction may not be accurate or robust relative to other models. While autoencoder models can be effective, they may not be the most appropriate choice for solving this problem.

TABLE III. result on training data

Metrics	**Deep Learning Models**		
	Neural Network	***1D CNN***	***Autoencoder***
Precision	0.93	0.93	0.82
Recall	0.82	0.88	0.82
F1 Score	0.87	0.90	0.82
Matthews Coef.	0.81	0.85	0.71
ROC AUC Score	0.93	0.93	0.85
Average Precision	0.93	0.90	0.74
Accuracy	0.91	0.93	0.86

VI. CONCLUSION

In this research, we study the use of Neural Networks, 1D CNN, and Autoencoders to predict whether the patient is dead or not. Our analysis shows that the deep learning 1D CNN model outperforms other models on various performance metrics such as Precision, Recall, F1-Score, Matthews Correlation Coefficient (MCC), ROC AUC Score, Average Precision, and Accuracy. Specifically, the 1D CNN model achieved a Precision of 0.93, Recall of 0.88, F1-Score of 0.90, MCC of 0.85, ROC AUC Score of 0.93, Average Precision of 0.90, and Accuracy of 0.93, outperforming the Neural Network and Autoencoder models in each of these metrics as shown in Table III. Using deep learning models with the 1D CNN architecture, physicians can improve their ability to predict patient outcomes and make more informed decisions about treatment strategies. Incorporating real-time patient data into the models can enhance predictive capabilities and enable more personalized health interventions. These findings may benefit healthcare by improving patient outcomes and the quality of care for cardiac patients.

REFERENCES

[1] T. Ahmad, A. Munir, S. Bhatti, M. Aftab, and M. Ali Raza, "Survival analysis of heart failure patients: A case study," PLoS One, vol. 12, p. e0181001, Jul. 2017, doi: 10.1371/journal.pone.0181001.

[2] M. Muntasir Nishat et al., "A Comprehensive Investigation of the Performances of Different Machine Learning Classifiers with SMOTE-ENN Oversampling Technique and Hyperparameter Optimization for Imbalanced Heart Failure Dataset," Sci Program, vol. 2022, 2022, doi: 10.1155/2022/3649406.

[3] N. D. WHO Team, "Cardiovascular diseases." Accessed: Jun. 10, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[4] V. K and S. Jyothi, "Decision Support System for Congenital Heart Disease Diagnosis based on Signs and Symptoms using Neural Networks," Int J Comput Appl, vol. 19, Apr. 2011, doi: 10.5120/2368-3115.

[5] A. J. Coats, "The pathophysiology of chronic heart failure.," Perfusion, vol. 15, no. 4, pp. 281–6, Jul. 2000, doi: 10.1177/026765910001500402.

[6] A. Newaz, N. Ahmed, and F. Haq, "Survival prediction of heart failure patients using machine learning techniques," Inform Med Unlocked, vol. 26, p. 100772, Jan. 2021, doi: 10.1016/j.imu.2021.100772.

[7] T. Ashine, G. Akessa, and K. Tadesse, "Assessing survival time of heart failure patients: using Bayesian approach," J Big Data, vol. 8, Dec. 2021, doi: 10.1186/s40537-021-00537-4.

[8] R. Katarya and P. Srinivas, Predicting Heart Disease at Early Stages using Machine Learning: A Survey. 2020. doi: 10.1109/ICESC48915.2020.9155586.

[9] R. Kasabe, "Heart Disease Prediction using Machine Learning," International Journal of Engineering Research and, vol. V9, Aug. 2020, doi: 10.17577/IJERTV9IS080128.

[10] C. Zheng et al., "Time-to-event prediction analysis of patients with chronic heart failure comorbid with atrial fibrillation: a LightGBM model," BMC Cardiovasc Disord, vol. 21, Aug. 2021, doi: 10.1186/s12872-021-02188-y.

[11] D. Chicco and G. Jurman, "Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone," BMC Med Inform Decis Mak, vol. 20, Feb. 2020, doi: 10.1186/s12911-020-1023-5.

[12] P. Moreno, "Improvement of a prediction model for heart failure survival through explainable artificial intelligence," Front Cardiovasc Med, vol. 10, Aug. 2023, doi: 10.3389/fcvm.2023.1219586.

[13] M. Tabassian et al., "Diagnosis of Heart Failure With Preserved Ejection Fraction: Machine Learning of Spatiotemporal Variations in Left Ventricular Deformation," Journal of the American Society of Echocardiography, vol. 31, Aug. 2018, doi: 10.1016/j.echo.2018.07.013.

[14] A. Tak, P. Parihar, S. Mathur, B. Das, and D. Sawal, Survival prediction in heart failure using machine learning algorithms. 2022. doi: 10.21203/rs.3.rs-1960150/v1.

[15] N. Basha, K. Gopal, and A. Kumar, "Early Detection of Heart Syndrome Using Machine Learning Technique," 2021.

[16] R. Najafi Vosough, J. Faradmal, K. Hosseini, A. Moghimbeigi, and H. Mahjub, "Predicting Hospital Readmission in Heart Failure Patients in Iran: A Comparison of Various Machine Learning Methods," Healthc Inform Res, vol. 27, pp. 307–314, Oct. 2021, doi: 10.4258/hir.2021.27.4.307.

[17] J.-M. Kwon et al., "Artificial intelligence algorithm for predicting mortality of patients with acute heart failure," PLoS One, vol. 14, p. e0219302, Jul. 2019, doi: 10.1371/journal.pone.0219302.

[18] C.-H. Lin, P.-K. Yang, Y.-C. Lin, and P.-K. Fu, On Machine Learning Models for Heart Disease Diagnosis. 2020. doi: 10.1109/ECBIOS50299.2020.9203614.

[19] T. Karayilan and O. Kilic, Prediction of heart disease using neural network. 2017. doi: 10.1109/UBMK.2017.8093512.

[20] S. Mishra, "A Comparative Study for Time-to-Event Analysis and Survival Prediction for Heart Failure Condition using Machine Learning Techniques," Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 4, pp. 115–134, Jul. 2022, doi: 10.35882/jeeemi.v4i3.225.

JETIR