Article **Journal of Physiotherapy and Medical Technology**
December, 2023 Vol.7 No.18 1-14

Heart failure prediction: Exploratory analysis and modeling with XGBoost and deep neural networks

Predicción de enfermedades cardíacas: análisis exploratorio y modelado con XGBoost y redes neuronales profundas

GUDIÑO-OCHOA, Alberto†*, OCHOA-ORNELAS, Raquel, URIBE-TOSCANO, Sofia and CUEVAS-CHÁVEZ, Jorge Ivan

Tecnológico Nacional de México - Instituto Tecnológico de Ciudad Guzmán, México.
Universidad de Guadalajara – Centro Universitario del Sur

ID 1st Author: *Alberto, Gudiño-Ochoa* / **ORC ID:** 0000-0002-2366-7452, **Researcher ID Thomson:** HDN-0235-2022, **arXiv Author ID:** AlbertoGO, **PubMed Author ID:** albertogudo, **CVU CONAHCYT ID:** 1135216

ID 1st Co-author: *Raquel, Ochoa-Ornelas* / **ORC ID:** 0000-0003-1824-5789, **Researcher ID Thomson:** S-4687-2018, **arXiv Author ID:** Raquel Ochoa, **SNI CONAHCYT ID:** 668976

ID 2nd Co-author: *Sofia, Uribe-Toscano* / **ORC ID:** 0009-0009-8325- 8237, **Researcher ID Thomson:** JRY-7988-2023, **arXiv Author ID:** Sofiuribest, **PubMed Author ID**: sofia.uribe9866

ID 3rd Co-author: *Jorge Ivan, Cuevas-Chávez* / **ORC ID:** 0009-0004-9377-0932, **Researcher ID Thomson:** JRX-8902-2023, **arXiv Author ID:** Jorge_Cuevas, **CONAHCYT:** 1220337

DOI: 10.35429/JP.2023.18.7.1.14 Received September 10, 2023; Accepted December 30, 2023

Abstract

Heart failure, a leading cause of global mortality, affects over 26 million individuals, contributing to approximately 17.9 million deaths annually representing 31% of all causes of death, as reported by the WHO. Early anticipation of this condition is paramount. Automated algorithms for early prediction of heart diseases, including machine learning models deployed in medical applications, have been extensively explored. This study scrutinizes data from 918 patients, some with a history of heart failure, drawn from diverse locations, encompassing Cleveland, Hungary, Switzerland, and Long Beach VA. The investigation centers on assessing predictive capability using XGBoost and deep neural networks in a sequential model. Results unveil promising precision—up to 88.04% and 88.58%, respectively—in forecasting future instances of heart diseases. Emphasizing the efficacy of these models, their potential as a valuable tool for medical professionals enabling early detection of this critical ailment is underscored. These findings underscore the significance of early prevention of heart diseases, potentially enhancing global health through advanced artificial intelligence techniques.

XGBoost, Heart disease prediction, Deep neural network

Resumen

La insuficiencia cardíaca, una de las principales causas de muerte a nivel mundial, afecta a más de 26 millones de personas y contribuye a aproximadamente 17.9 millones de fallecimientos anuales, representando el 31% de todas las causas de muerte según la OMS. La anticipación temprana de esta condición es vital. Se han explorado algoritmos automatizados para la predicción temprana de enfermedades cardíacas, incluyendo modelos de aprendizaje automático implementados en aplicaciones médicas. Este estudio analiza datos de 918 pacientes, algunos con historial de insuficiencia cardíaca, provenientes de observaciones en diferentes localidades, incluyendo Cleveland, Hungría, Suiza y Long Beach VA. La investigación se enfoca en evaluar la capacidad predictiva mediante el uso de XGBoost y redes neuronales profundas en un modelo secuencial. Los resultados revelan una precisión prometedora: hasta un 88.04% y 88.58%, respectivamente, en la predicción de futuros casos de enfermedades cardíacas. Destacando la efectividad de estos modelos, se plantea su potencial como herramienta valiosa para los profesionales médicos, permitiendo la detección precoz de esta enfermedad crítica. Estos hallazgos resaltan la importancia de la prevención temprana de enfermedades cardíacas para mejorar significativamente la salud a nivel global a tráves de técnicas avanzadas de inteligencia artificial.

XGBoost, Predicción de enfermedad cardiaca, Red neuronal profunda

Citation: GUDIÑO-OCHOA, Alberto, OCHOA-ORNELAS, Raquel, URIBE-TOSCANO, Sofia and CUEVAS-CHÁVEZ, Jorge Ivan. Heart failure prediction: Exploratory analysis and modeling with XGBoost and deep neural networks Journal of Physiotherapy and Medical Technology. 2023. 7-18: 1-14

* Correspondence to the Author (e-mail: albertogudo@hotmail.com)
† Researcher contributing as first author

Introduction

In the expansive realm of medical data and the burgeoning field of data science and analysis, innovative efforts by startups have undertaken the formidable task of creating predictive markers for impending diseases (Leung et al., 2020). Among these conditions, cardiovascular diseases (CVDs) stand out, claiming a staggering 17.9 million lives annually, representing a significant 31% of global mortality (Levy et al., 2017). Within this landscape, heart failure emerges as a prevalent consequence of CVDs (Reddy et al., 2021).

Individuals grappling with cardiovascular ailments or teetering on the edge of escalated cardiovascular risk, often linked to factors like hypertension, diabetes, hyperlipidemia, or existing medical conditions, require early identification and adept management (Kullo et al., 2010). Herein lies the potential of machine learning models to offer their guidance (Louridi et al., 2021). This technological synergy not only seeks to automate a natural conundrum but also to illuminate forthcoming challenges, all driven by the formidable capacities of artificial intelligence techniques (Lv et al., 2021; Chang et al., 2022).

Previous research showcases the integration of machine learning models with medical information systems to predict heart failure or other diseases using patient-collected data. Effective models based on decision trees, Naive Bayes, Random Forest, logistic regression, and Support Vector Machines (SVM) have achieved a maximum accuracy of 88% (Alotaibi, 2019; Mansur Huang et al., 2021; Kedia et al., 2022). Data normalization and preprocessing are pivotal for enhancing model accuracy (Wang, 2021). Hyperparameter optimization techniques such as TPOT and Random Forest have achieved the highest accuracy of 97.52% (Valarmathi et al., 2021). The application of deep neural networks, such as Recurrent Neural Networks (RNN), has been instrumental in predicting the risk and onset of heart failure, achieving an 82% accuracy, demonstrating their efficacy in predictive modeling using electronic medical records (Rasmy et al., 2018). Even models based on Multilayer Perceptron have reached an 88% accuracy (Awan et al., 2019; Lee et al., 2020).

People with cardiovascular diseases or those at high cardiovascular risk (due to factors such as hypertension, diabetes, hyperlipidemia, or existing diseases) require early detection and management, wherein machine learning models play a crucial role. Moreover, with environmental deterioration, these factors may lead to an increase in heart failure cases in the future. Neglecting the issue of heart failure could inevitably result in fatalities (Kim et al., 2021).

Four out of 5 CVD-related deaths emanate from heart attacks and strokes, with a third occurring prematurely among individuals under 70 years (Jagannathan et al., 2019). Heart failure, often an outcome of CVDs, serves as a focal point, addressed through a dataset encompassing 11 features pivotal in predicting cardiac diseases. For the current study, a dataset amalgamating five previously independent datasets: Cleveland (303 observations), Hungarian (294 observations), Swiss (123 observations), Long Beach VA (200 observations), and Stalog (270 observations), aggregated to a total of 1190 observations, identified 272 instances of duplications, culminating in a refined set of 918 observations.

This comprehensive dataset, available on Kaggle (fedesoriano, 2021), stands as the most extensive compilation in cardiovascular disease research to date. Table 1 presents a description of the pathological information collected. Among the most notable data from studies with a similar focus are 5 risk factors (body mass index, systolic blood pressure, high-density lipoprotein cholesterol, current smoking, and diabetes) and the incidence of cardiovascular diseases and death from any cause through Cox regression analysis (Global Cardiovascular Risk Consortium, 2023).; this study contemplates three of these primary factors.

This research's focus lies in implementing an Exploratory Data Analysis (EDA) to comprehend the dataset, preprocessing pathological information through standardization, and proceeding with the implementation of advanced machine learning techniques via XGBoost. The optimization of hyperparameters using the GridSearch method will be followed by evaluating prediction metrics with patients' pathological information, presenting an acceptable model for the future prediction of patients who might develop early-stage cardiac diseases.

Additionally, a comparison and evaluation of a Deep Neural Network (DNN) optimized to ascertain its accuracy in future classifications will provide valuable insights into AI-based predictive models for medical diagnoses.

Parameter	Description	Type
Age	Years	Patient`s age
Sex	patient's sex	M: Male, F: Female
Chest Pain Type	type of chest pain	TA: Typical Angina, ATA: Atypical Angina, NAP: Non-Anginal Pain, ASY: Asymptomatic
Resting Blood Pressure	resting blood pressure	mmHg
Cholesterol	serum cholesterol	mm/dl
Fasting Blood Sugar	fasting blood sugar level	1: if Fasting Blood Sugar > 120 mg/dl, 0: otherwise
Resting ECG	resting electrocardiogram results	Normal: Normal, ST: ST-T wave abnormality (T wave inversions and/or ST elevation or depression > 0.05 mV), LVH: probable or definite left ventricular hypertrophy by Estes' criteria
Maximum Heart Rate	maximum heart rate achieved	Numeric value between 60 and 202
Exercise-Induced Angina	exercise-induced angina	Y: Yes, N: No
Oldpeak	oldpeak = ST	Numeric value measured in depression
ST Slope	the slope of the peak exercise ST segment	Up: upsloping, Flat: flat, down: downsloping
Heart Disease	output class	1: heart disease, 0: Normal

Table 1 Information and description of the dataset.

XGBoost in Predictions of Cardiac Diseases

XGBoost, a machine learning technique, harnesses the ensemble method to merge multiple weak decision tree models, crafting a more robust and accurate model. Grounded in decision trees, known for their shallow depth and slightly superior performance to random predictions, XGBoost employs boosting to sequentially train trees, each rectifying the errors of the previous model (Chen et al., 2015).

Optimization in XGBoost revolves around an objective function balancing model loss and complexity. This technique utilizes a loss function to evaluate prediction accuracy, offering advantages like high performance, computational efficiency, and control over overfitting through regularization (Nielsen et al., 2016). For log-loss classification, the objective function is expressed by equation (1), where various parameters assess model predictions and regularization hyperparameters (Nielsen et al., 2016).

$$\text{Classificaction} = \sum_{\text{i}=1}^{\text{n}} (y_i \log(p_i) + (1 - y_i)\log(1 - p_i)) + \gamma \text{T} + \frac{1}{2}\lambda \sum_{j=1}^{T} \omega_j^2 \quad (1)$$

where: y_i represents the actual labels, p_i or y_i denote the model predictions, T stands for the number of tree leaves, ω_j^2 signifies the square of the weight assigned to each tree leaf, and λ as well as γ represent regularization hyperparameters.

In the classification of cardiac diseases, XGBoost has exhibited high performance in various scoring metrics using logistic regression (Nalluri et al., 2020). With optimized hyperparameters using OPTUNA, it achieved accuracies of 94.7%, 89.3%, and 88.5% in classifying distinct cardiac disease datasets (Srinivas et al., 2022). Additionally, Bayesian optimization combined with encoding techniques reached 91.8% accuracy in predicting heart disease in clinical settings (Budholiya et al., 2022). Utilizing feature selection via information gain and a hybrid Smote-Enn algorithm for imbalanced datasets in cardiac disease prediction attained 93.44% accuracy with XGBoost (Yang et al., 2022). Moreover, this model has demonstrated reliability in assessing cardiovascular disease risk in patients with type 2 diabetes mellitus (Athanasiou et al., 2020).

Deep Neural Networks (DNNs)

Deep or Dense Neural Networks (DNNs) consist of interconnected layers, where each neuron in one layer connects to all neurons in the subsequent layer, creating a densely connected network. Apart from input and output layers, hidden layers perform data transformation operations; activation functions are derived from neurons in a hidden layer by considering weighted inputs from neurons in preceding layers (Nazari et al., 2021).

Common examples of activation functions include Rectified Linear Unit (ReLU), sigmoid, hyperbolic tangent, among others (Ding et al., 2018). This algorithm is considered a deep and automatic learning method, adjusting neuron connection weights and biases during training by minimizing loss functions. Regularization techniques like Dropout, L1, L2 are employed in DNNs to prevent overfitting and enhance generalization (Cogswell et al., 2015). The key formula is presented in the following equation (2).

$$\text{Output} = \sigma\left(\sum_{i=1}^{n} \omega_i \cdot x_i + b\right) \quad (2)$$

where: x_i represents inputs, ω_i denotes the weights associated with each input, b stands for bias, and σ represents the activation function. DNNs possess the capability to learn intricate relationships within nonlinear data and automatically extract features (Zeng et al., 2014). Achieving up to 90% accuracy in predicting cardiac diseases has been accomplished using an enhanced Sparse Autoencoder (SAE) integrated with neural networks (Mienye et al., 2020). Assessment of metrics across epochs during DNN training has demonstrated a classification accuracy ranging between 80-90% (Ramprakash et al., 2020).

Exploratory Data Analysis (EDA)

During the phase of understanding the dataset, an Exploratory Data Analysis (EDA) was performed using Python-based tools. This aimed to uncover trends, relationships, and distributions within the data through statistical methods. Specifically, it delved into patient information, distinguishing between individuals with cardiac ailments and those deemed healthy. Additionally, it sought to offer a medical perspective by examining pertinent characteristics within patients' pathological data, such as clinical studies, age, and gender.

Among the 918 patients in the dataset, 508 were diagnosed with heart failure, while 410 did not exhibit such conditions. Figure 1 depicts bar graphs showcasing a nearly balanced distribution between healthy patients and those affected by cardiac disease. The median age of patients diagnosed with heart disease was 57, whereas individuals without heart disease showed a slightly younger median age of 51.

Previous similar clinical analyses found a comparable average age of 54.4 years (Global Cardiovascular Risk Consortium, 2023). The breakdown is nearly 90.2% male patients and 9.8% female patients in the dataset. Figure 2 showcases the prevalence of cardiac disease among these genders. Remarkably, among the 725 male patients, 63.2% have been diagnosed with cardiac disease. In contrast, among females, the diagnosis reveals that one in every four females has been diagnosed with a cardiac condition.

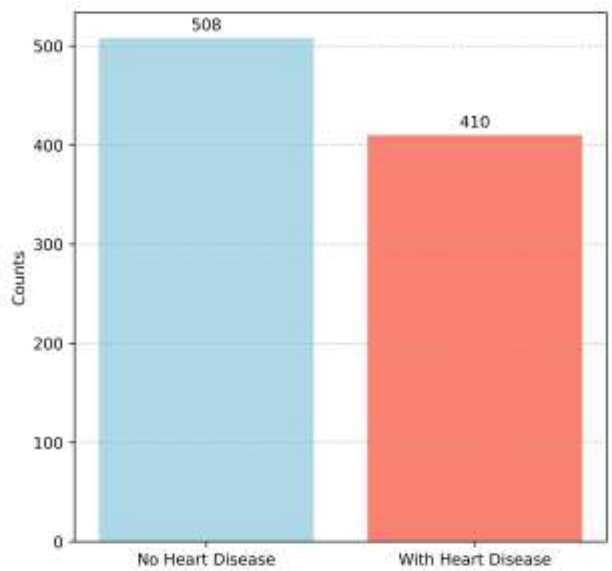

Figure 1 Total number of healthy patients and patients with heart disease.

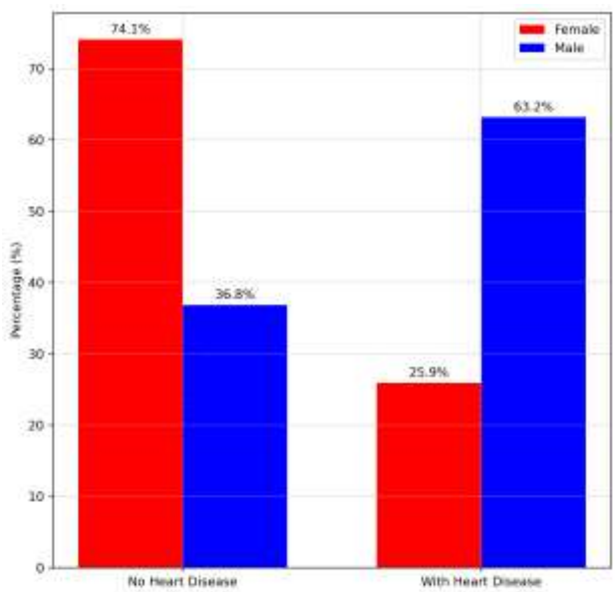

Figure 2 Percentage of patients and prevalence of heart disease among men and women.

Figure 3 displays boxplots depicting distinct age distributions among patients with and without heart disease.

ISSN 2523-6849

December, 2023 Vol.7 No.18 1-14

Among those diagnosed with heart disease, the boxplot indicates a narrower age range, mainly spanning from 51 to 62 years, with most patients falling within this interval. Notably, there are a few younger outliers below the lower whisker in this group. In contrast, patients without heart disease show a slightly broader age spectrum, evenly distributed without outliers. Many individuals in this category fall within the age range of 43 to 57 years, representing a relatively younger population.

Additionally, within Figure 3, the boxplots for Systolic Blood Pressure exhibit similar distributions between the groups. Both sets show upper and lower outliers, with most patients' blood pressure ranging from 120 to 145 mmHg. The median blood pressure remains consistent at approximately 130 mmHg for both patient cohorts. These data reveal the relationship highlighted in previous medical studies, linking severe increases in blood pressure directly to higher risks of developing heart failure.

The progression from hypertension to heart failure is complex and multifaceted. Therefore, while it elevates the risk of heart failure, not everyone with hypertension develops it. It would be challenging to determine heart failure solely based on hypertension, but it significantly increases the risk (Di Palo et al., 2020).

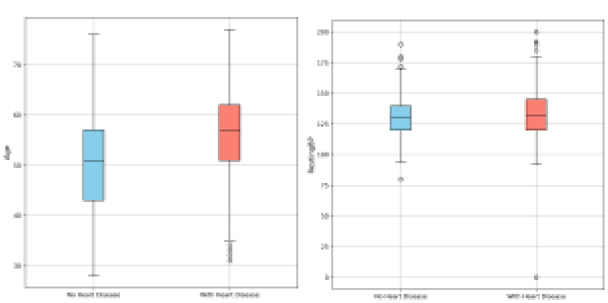

Figure 3 Boxplots depicting age distributions (left) and Systolic Blood Pressure (right).

In Figure 4, the distribution of cholesterol levels appears right-skewed, particularly within patients with cardiac conditions where numerous observations show missing cholesterol levels, marked as 0. On the other hand, among healthy patients, a median of 225 is evident within a range spanning approximately 200 to 275, displaying several outliers, as depicted in the boxplots.

However, it is well-established that lower levels of cholesterol could be detrimental to patients with Heart Failure. This could happen due to treatment for hypercholesterolemia with statins, which might pose risks for Heart Failure patients as they decrease Coenzyme Q10 concentrations. This concentration is vital in the respiratory reactions of cardiac myofibrils and mitochondria. Hence, it's not uncommon to find atypical or very low cholesterol levels in patients with heart failure (Charach et al., 2023).

Also, within the same Figure, it's noticeable that patients without heart disease exhibit the potential for higher maximum heart rates compared to patients with heart disease. The median heart rate among patients without heart disease stands at 150 beats per minute, contrasting with a median of 126 beats per minute among those with heart disease. Hypercholesterolemia is the primary risk factor leading to ischemic heart disease (a significant cause of heart failure). Chronotropic incompetence, a phenomenon common in heart failure with preserved ejection fraction (HFpEF), is linked to impaired aerobic capacity. Consequently, patients with heart failure struggle to achieve a higher maximum heart rate (Sarma et al., 2020).

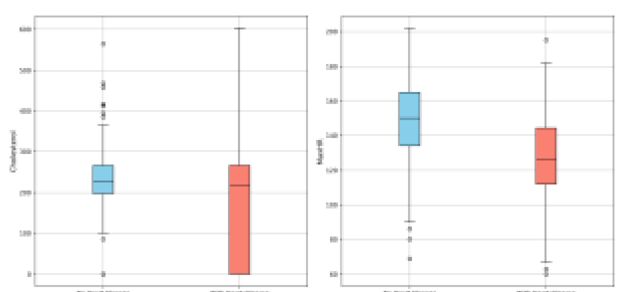

Figure 4 Boxplots illustrating cholesterol distribution (left) and maximum heart rate achieved (right).

In Figure 5, the boxplots distinctly illustrate the disparity in ST segment depression distribution between the two patient cohorts. Among those diagnosed with heart disease, there's a noticeably wider variability in ST depression, with several markedly larger outliers. Most patients in this category exhibit ST depressions ranging from 0 to 2 mm, centered around a median value of 1.2 mm.

Clinical investigations have demonstrated the association of ST segment depression in the initial ECG during atrial fibrillation with clinical outcomes.

ISSN 2523-6849

The primary endpoint was heart failure: cardiac death or hospitalization due to heart failure. The prevalence of ST segment depression was 25.4%; the incidence rate of the composite endpoint of heart failure was significantly higher in patients with ST segment depression compared to those without (5.3% vs. 3.6% per patient-year, logarithmic range $P < 0.01$). Additionally, ST segment depression was an independent predictor for the composite endpoint of heart failure (hazard ratio 1.23, 95% confidence interval: 1.03–1.49, $P = 0.03$) (Kawaji et al., 2023).

The patients within the heart disease subset of the Kaggle dataset seem to fall within this range, suggesting a potential clinical relevance like the relationship between ST segment depression and heart failure outcomes.

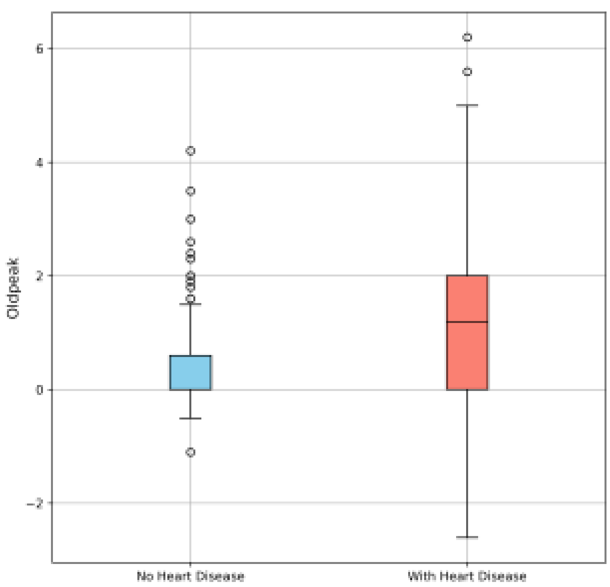

Figure 5 Boxplots illustrating oldpeak ST among individuals with and without heart disease.

Conversely, among patients without heart disease, the range of ST depression is comparatively narrower, spanning from 0 to 0.6 mm, with a median ST depression value of 0 mm. However, it's important to note that the distribution in this group displays a higher degree of skewness overall.

In Figure 6, it's evident that among patients diagnosed with diabetes, approximately 80% exhibit concurrent heart disease. This observation is grounded in the inherent relationship between heart failure and diabetes mellitus, representing substantial morbidity and mortality worldwide.

Furthermore, these chronic conditions can coexist and influence each other: among diabetic adults (> 64 years), a reported prevalence of heart failure stands at 22%. Patients with heart failure carry an increased risk of developing diabetes mellitus, and conversely, patients with diabetes mellitus have a higher likelihood of developing heart failure (Kallas et al., 2023).

Additionally, the graphic on the left side illustrates a notably heightened prevalence of heart disease, exceeding 85%, among patients experiencing exercise-induced angina. However, exercise-induced vasospastic angina represents a relatively uncommon clinical scenario characterized by the onset of chest symptoms during exertion due to coronary vasospasm. Unless a positive provocation test is achieved, diagnosing this can be challenging, potentially resulting in adverse cardiac events, including myocardial infarction, ventricular arrhythmia, and sudden cardiac arrest. A definitive diagnosis of vasospastic angina often poses a challenge because these spasms tend to be transient, and many coronary spasm events are asymptomatic (Tamura et al., 2018).

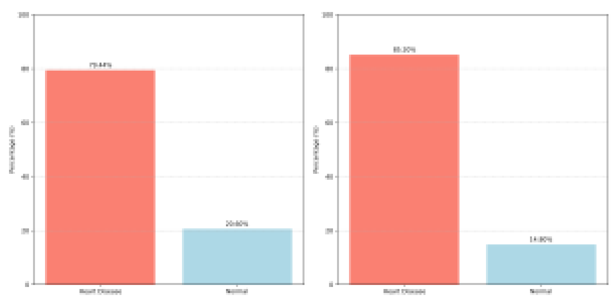

Figure 6 Prevalence depicted as percentages of patients with heart disease and healthy individuals among those with diabetes (left) and exercise-induced angina (right).

Exploring the categories of characteristics reveals a variability in the types of resting electrocardiogram results, as depicted in Figure 7. The distribution is as follows: 552 patients with normal ECG, accounting for 56.1% of the sample; 178 patients exhibiting ST-T wave abnormalities, representing 23%; and 188 patients showing probable or definite left ventricular hypertrophy according to Estes' criteria, making up 20.9% of the dataset.

Article Journal of Physiotherapy and Medical Technology
December, 2023 Vol.7 No.18 1-14

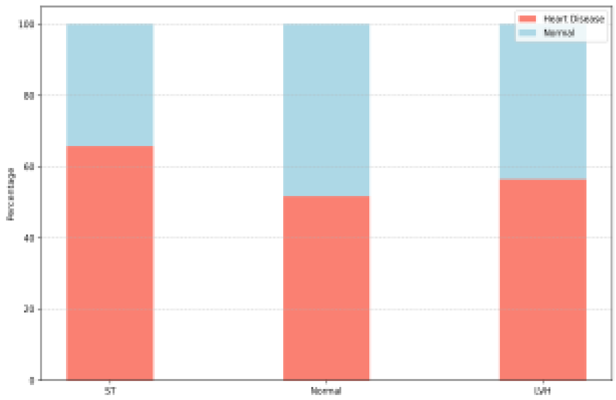

Figure 7 Prevalence of heart disease categorized by resting ECG.

Analyzing the resting ECG test outcomes reveals that over 65% of patients diagnosed with heart disease exhibit ST-T wave abnormalities in their ECG readings, marking the highest proportion among the different groups.

Studies focusing on resting electrocardiographic results have shown similarities in ST-T wave abnormalities and probable or definite left ventricular hypertrophy according to the Romhilt-Estes criteria (Aghamohammadi et al., 2019).

Three-quarters (75%) of the dataset, totaling 460 patients, exhibit a flat slope of the peak exercise ST segment. Additionally, 395 patients (15.4%) showcase an upward ST segment slope, while 63 patients (2.6%) display a downward ST segment slope. Remarkably, patients with a flat or downward ST segment slope during exercise present the highest prevalence of cardiovascular disease at 82.8% and 77.8%, respectively, as shown in Figure 8 (Aghamohammadi et al., 2019).

This relationship is similar to the findings where patients with ST segment depression were older and had more comorbidities than those without. The higher risk was observed in the horizontal or downward ST segment depression, but not in the upward depression. Moreover, among 40 patients exhibiting this pattern, 19 are at risk of a heart attack, with 3 of these 19 patients being women and 16 men. Out of these, 16 patients experience level 4 chest pain (Asymptomatic), 2 of whom have atypical angina, and only one has typical angina. Hence, it's notable that asymptomatic chest pain is also more prevalent in this research.

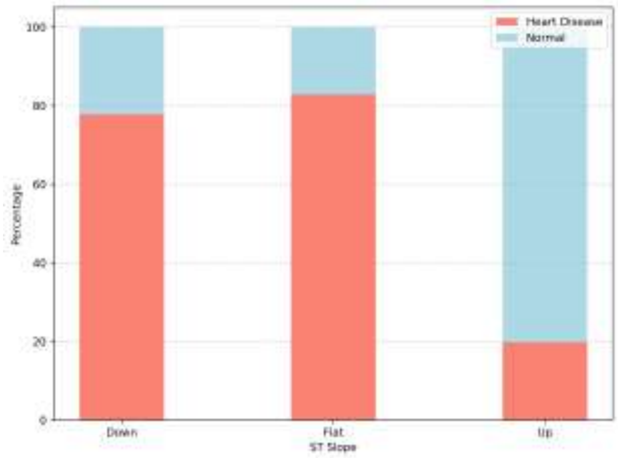

Figure 8 Prevalence of heart disease categorized by ST Slope.

The predominant type of chest pain noted is asymptomatic, observed in a total of 496 patients, constituting 77.2% of cases. Following this, non-anginal pain is reported in 203 patients, accounting for 14.2%, with atypical angina observed in 173 patients, representing 4.7% of the cohort. Lastly, 46 patients present with typical angina, comprising 3.9% of the dataset.

Figure 9 demonstrates that among individuals diagnosed with heart disease, asymptomatic chest pain prevails, exceeding 77%. Additionally, the incidence of heart disease is notably higher in men compared to women, occurring at a ratio of nearly 9 to 1 among patients diagnosed with cardiovascular conditions.

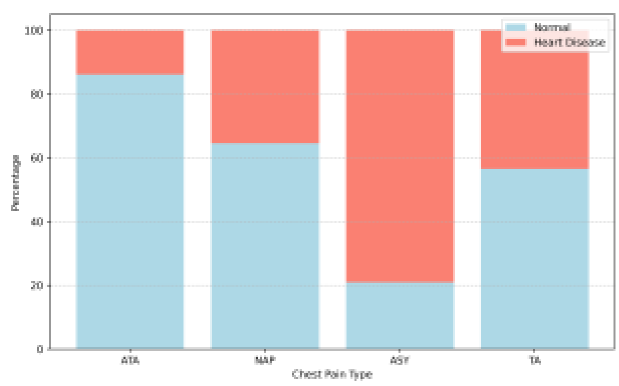

Figure 9 Prevalence of chest pain among individuals diagnosed with heart disease and those without heart issues.

Figure 10 illustrates the correlations and scatterplots, revealing significant associations within the dataset. Heart disease exhibits the strongest positive correlation with OldPeak (correlation = 0.4) and the most robust negative correlation with MaxHR (correlation = -0.4).

ISSN 2523-6849

Additionally, a moderately strong relationship of -0.38 exists between Age and MaxHR, indicating that as age increases, heart rate tends to decrease.

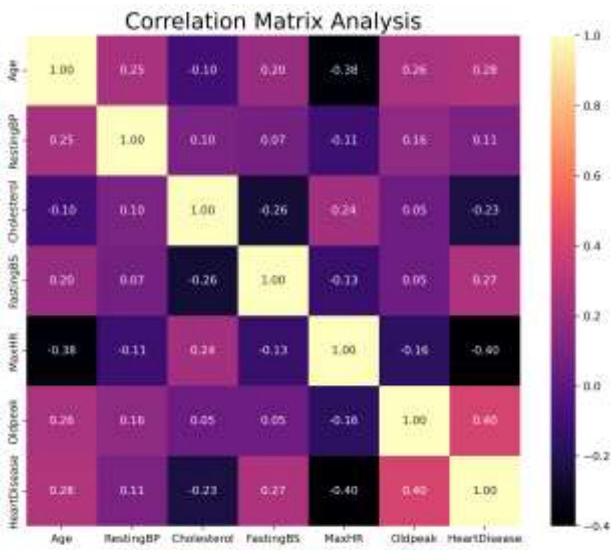

Figure 10 Correlation Matrix Analysis

Data Preprocessing and Classification Parameters with XGBoost and Deep Neural Networks

Preprocessing stands as a fundamental phase preceding model training. Within this study, numerical attributes undergo standardization, involving the removal of mean and scaling to unit variance. Meanwhile, categorical attributes are subjected to one-hot encoding, essential for the accurate interpretation of categorical data by the machine learning model.

Standardization, although not universally mandated, is widely considered a best practice in model training. Its application ensures uniformity in feature scaling, aiding model convergence and performance. On the other hand, one-hot encoding transforms categorical variables into a binary format, allowing the model to effectively comprehend categorical information.

The implementation of StandardScaler in sklearn operates under the assumption that the data, represented by variable, might not adhere strictly to a Gaussian distribution. However, it orchestrates transformations to align the distribution to a mean value of 0 and a standard deviation of 1. In other words, given a feature vector x, it modifies the values as follows:

$$\mathrm{Y_i} = \frac{x_i - \mu(\vec{x})}{\sigma(\vec{x})} \quad (1)$$

where: x_i is the i-th element of the original feature vector $(\vec{x})$, $\mu(\vec{x})$ is the mean of the feature vector, and $\sigma(\vec{x})$ is the standard deviation of the feature vector.

Results and discussion

The XGBoost model was trained using 70% of preprocessed data as the training set (642 patients) and the remaining 30% as the test set (276 patients). The GridSearchCV technique was employed to fine-tune hyperparameters, focusing on binary logistic classification to distinguish patients prone to heart disease from healthy individuals.

Optimal hyperparameters for the XGBoost algorithm are presented in Table 2, highlighting the use of regularization via alpha and lambda to prevent overfitting during data classification.

Noteworthy are the performance metrics obtained from the classification report, demonstrating an 88.04% precision in the test set. Additionally, a 90% recall was achieved in identifying patients with heart disease. Table 3 details precision, recall, f1-score, and support values for the binary classification.

Parameter	Value	Description
Learning_rate	0.05	Controls the contribution of each tree to the model.
n_estimators	200	Specifies the number of trees used in the model.
objective	'binary: logistic'	binary classification
max_depth	3	Maximum depth of each tree, limiting the number of splits in each tree.
reg_alpha	1	avoid overfitting
reg_lambda	0.1	control overfitting

Table 2 XGBoost best hyperparameters with Gridsearch**.**

Condition	Precision	Recall	F1-Score	Support
0: Normal	0.88	0.85	0.86	123
1: Heart disease	0.88	0.90	0.89	153

Table 3 Classification report with XGBoost-GridSearchCV

The confusion matrix in Figure 11 illustrates 105 accurate predictions for healthy patients and 138 correct predictions for those diagnosed with heart disease.

However, there are 18 false negatives where healthy patients were misclassified as having a heart condition, along with 15 false positives, indicating patients with heart disease incorrectly classified as healthy or normal.

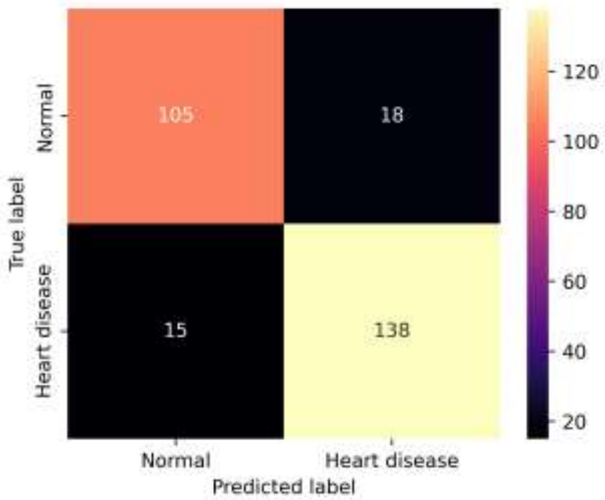

Figure 11 Confusion Matrix depicting results from the XGBoost classifier.

Furthermore, the implemented DNN model is based on a sequential model imported from the Keras library. The test set was split into 20%, with the remaining 80% designated for training. Within this 80%, an additional 25% was set aside to form the validation set. Thus, 60% of the total data is allocated to training, 20% to testing, and the remaining 20% to validation. Table 4 details the parameters and characteristics of the neural network.

Parameter	Value	Description
Layers (Dense)	Four layers, one serving as input and one as output.	Controlling the contribution of each layer to the model's predictive capacity.
Activation fucntion	'ReLU' is applied in the model's intermediate layers, while 'sigmoid' is used in the final layer.	'Sigmoid' is commonly employed in binary classification tasks to obtain probabilities between 0 and 1.
Regularization (L1)	0.001 for second layer and 0.0001 for third layer of dense neurons.	Aimed at constraining the weights to mitigate overfitting.
Dropout	0.5, 0.5, 0.25	Dropout layers are employed after each dense layer with dropout rates set at 0.5, 0.5, and 0.25, sequentially, to alleviate overfitting by randomly deactivating a fraction of neurons during training.

Table 4 Architecture of DNN

The optimizer employed during training was 'adam' due to its stochastic gradient descent characteristics. Binary cross-entropy loss function was utilized, deemed suitable for this classification problem. The evaluation metric utilized to assess the model during training was 'accuracy'.

The model was trained on the training data for 300 epochs with a batch size of 32. Simultaneously, validation was conducted during training to monitor performance and prevent overfitting. Figure 12 illustrates the model's performance, showcasing a slight discrepancy between the error rates of the training and validation sets.

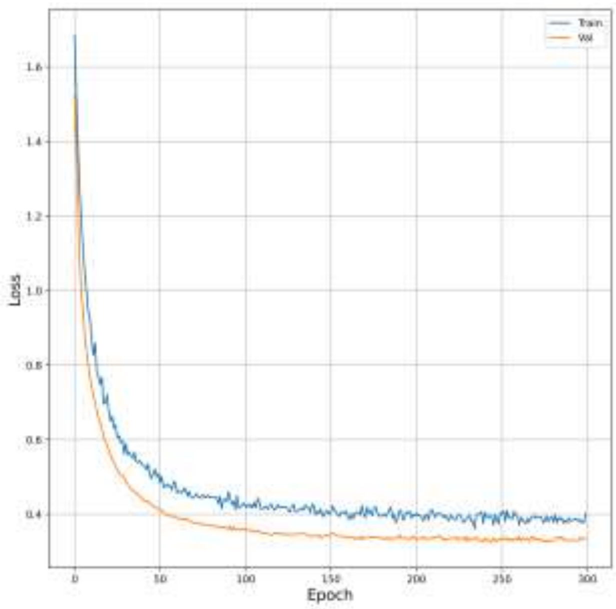

Figure 12 DNN Model Loss with Training and Validation Data.

The evaluation displayed an 88.58% accuracy across both datasets, as depicted in Figure 13. Detailed scoring metrics can be found in the classification report provided in Table 5. Comparatively, the DNN model surpasses XGBoost in terms of recall and f1-score, achieving an 88% and 87% detection rate for healthy patients, respectively. Moreover, it showcases significant precision at 91% for patients with heart disease and achieves an f1-score of 90%.

Article **Journal of Physiotherapy and Medical Technology**
December, 2023 Vol.7 No.18 1-14

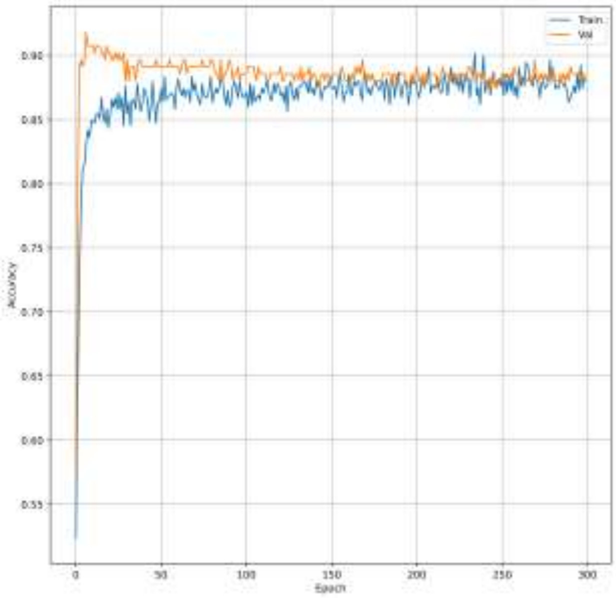

Figure 13 Model Accuracy During Training with Training and Validation Data

Condition	Precision	Recall	F1-Score	Support
0: Normal	0.85	0.88	0.87	77
1: Heart disease	0.91	0.89	0.90	153

Table 5 Classification Report for the DNN Model

The classification using DNN demonstrates a slight superiority compared to the XGBoost model. In this case, Figure 14 shows 68 accurate predictions for healthy patients and 95 correct predictions for those diagnosed with heart disease. However, only 12 false positives and 9 false negatives were detected.

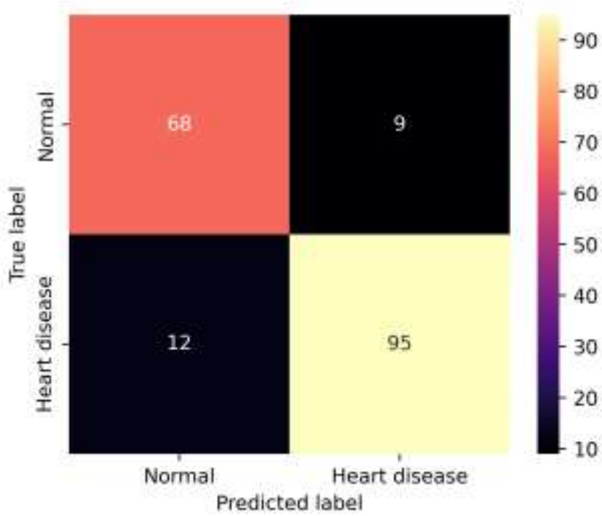

Figure 14 Confusion Matrix illustrating results from the DNN Classifier

The ROC curve metrics in Figure 15 demonstrate a value of 0.885 for the DNN and 0.878 for the XGBoost, with the DNN having the value closest to the true positive prediction.

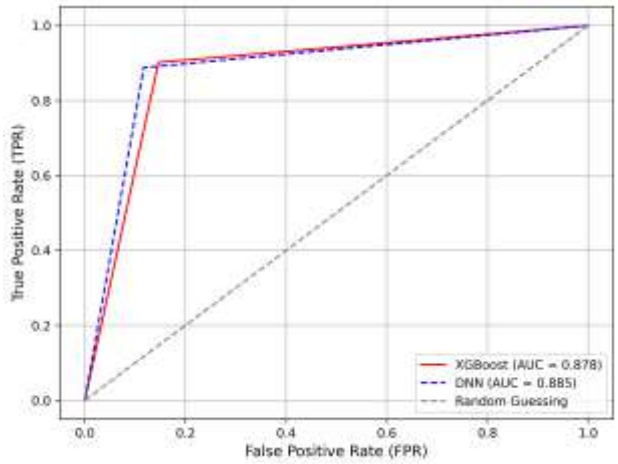

Figure 15 ROC Curves Comparison between the DNN and XGBoost Models.

The obtained results showcase significant metrics in Figure 16 concerning the detection of cardiac diseases. The XGBoost model demonstrates superior recall performance at 90.2% compared to the DNN's 88.79%. This signifies the XGBoost's efficacy in capturing most actual cardiac disease cases. Specifically, within this balanced patient dataset, the preprocessing and classification techniques employed by both XGBoost and DNN are on par or even surpassing other methodologies (Srinivas et al., 2022). Missing a positive case could have severe health consequences for the patient, underscoring the criticality of maximizing recall. Notably, the precision remains similar between both models at 88%.

This precision is crucial to prevent misdiagnosis, averting unnecessary treatments or undue stress for the patient. Finally, the DNN achieves a superior F1-score of 90.05%. As a harmonic mean between precision and recall, it emphasizes the precise identification of positive cases and the minimization of false positives compared to XGBoost.

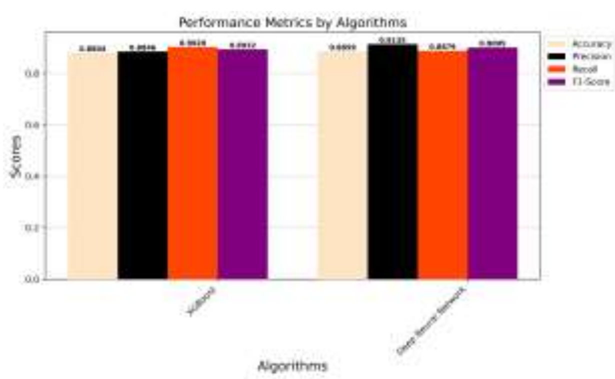

Figure 16 Performance Metrics by Algorithms between the DNN and XGBoost Models

ISSN 2523-6849

Conclusions

The proposed models utilizing machine learning classifiers for cardiac disease detection were rigorously compared between DNN and XGBoost, leveraging the best hyperparameters obtained through Gridsearch. While both models demonstrated comparable precision, XGBoost exhibited superior recall at 90.2%, whereas DNN showcased stronger classification and F1-score metrics at 90%. These findings underscore the potential of machine learning in refining disease detection, especially in a cardiac context.

This study benefited significantly from a specific dataset, unveiling critical medical insights such as the correlation between certain parameters—like diabetes—and increased risk of cardiac illness. Moreover, the prevalence of chest pain, particularly among asymptomatic patients and those experiencing exercise-induced angina, proved crucial in identifying potential cardiac issues. The methodology's strength lay in meticulous data preprocessing, where numeric values and categorical classifications based on clinical and pathological test types were pivotal in enhancing both XGBoost and DNN classifications.

Looking ahead, this research paves the way for further investigations, potentially integrating additional datasets or exploring hybrid models amalgamating the strengths of various machine learning techniques. Moreover, the models' potential generalizability beyond this specific dataset merits consideration, exploring adaptations required for diverse populations.

The findings presented here hold promise for revolutionizing early detection strategies for cardiac diseases, offering profound implications for improved patient care and healthcare management. This work contributes to advancing the field, introducing novel methodologies and insights that may drive future breakthroughs in cardiac disease detection and management.

Acknowledgments

The authors would like to extend our heartfelt gratitude to the authors whose invaluable contributions facilitated this research endeavor through the provision of the Heart Failure Prediction Dataset. Specifically, we extend our thanks to the Hungarian Institute of Cardiology in Budapest, represented by Andras Janosi, M.D. Our appreciation also goes to the University Hospital in Zurich, Switzerland, and the esteemed William Steinbrunn, M.D., as well as the University Hospital in Basel, Switzerland, with gratitude to Matthias Pfisterer, M.D. Additionally, our sincere thanks to the V.A. Medical Center in Long Beach and the Cleveland Clinic Foundation, represented by Robert Detrano, M.D., Ph.D.

Their dedication and efforts in curating and making available this dataset have been instrumental in advancing scientific research in the field of heart failure prediction.

References

Alotaibi, F. S. (2019). Implementation of machine learning model to predict heart failure disease. International Journal of Advanced Computer Science and Applications, 10(6).

Athanasiou, M., Sfrintzeri, K., Zarkogianni, K., Thanopoulou, A. C., & Nikita, K. S. (2020, October). An explainable XGBoost–based approach towards assessing the risk of cardiovascular disease in patients with Type 2 Diabetes Mellitus. In 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE) (pp. 859-864). IEEE. URL: https://ieeexplore.ieee.org/abstract/document/9288053. DOI: 10.1109/BIBE50027.2020.00146.

Awan, S. E., Bennamoun, M., Sohel, F., Sanfilippo, F. M., & Dwivedi, G. (2019). Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. ESC heart failure, 6(2), 428-435. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/ehf2.12419. DOI: 10.1002/ehf2.12419.

Budholiya, K., Shrivastava, S. K., & Sharma, V. (2022). An optimized XGBoost based diagnostic system for effective prediction of heart disease. Journal of King Saud University-Computer and Information Sciences, 34(7), 4514-4523. URL: https://www.sciencedirect.com/science/article/pii/S1319157820304936. DOI: 10.1016/j.jksuci.2020.10.013

Chang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. A. (2022). An artificial intelligence model for heart disease detection using machine learning algorithms. Healthcare Analytics, 2, 100016. URL: https://www.sciencedirect.com/science/article/pii/S2772442522000016. DOI: 10.1016/j.health.2022.100016.

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., ... & Zhou, T. (2015). Xgboost: extreme gradient boosting. R package version 0.4-2, 1(4), 1-4. URL: https://cran.ms.unimelb.edu.au/web/packages/xgboost/vignettes/xgboost.pdf.

Charach, L., Grosskopf, I., Galin, L., Guterman, I., Karniel, E., & Charach, G. (2023). Low Cholesterol Levels in Younger Heart Failure Patients May Predict Unfavorable Outcomes. Medicina, 59(7), 1314. URL: https://www.mdpi.com/1648-9144/59/7/1314. DOI: 10.3390/medicina59071314.

Cogswell, M., Ahmed, F., Girshick, R., Zitnick, L., & Batra, D. (2015). Reducing overfitting in deep networks by decorrelating representations. arXiv preprint arXiv:1511.06068. URL: https://arxiv.org/abs/1511.06068. DOI: 10.48550/arXiv.1511.06068.

Di Palo, K. E., & Barone, N. J. (2020). Hypertension and heart failure: prevention, targets, and treatment. Heart failure clinics, 16(1), 99-106. URL: https://www.heartfailure.theclinics.com/article/S1551-7136(19)30099-6/fulltext. DOI: 10.1016/j.hfc.2019.09.001.

Ding, B., Qian, H., & Zhou, J. (2018, June). Activation functions and their characteristics in deep neural networks. In 2018 Chinese control and decision conference (CCDC) (pp. 1836-1841). IEEE. URL: https://ieeexplore.ieee.org/abstract/document/8407425. DOI: 10.1109/CCDC.2018.8407425.

fedesoriano. (September 2021). Heart Failure Prediction Dataset. Retrieved September 2023 from https://www.kaggle.com/fedesoriano/heart-failure-prediction.

Global Cardiovascular Risk Consortium. (2023). Global effect of modifiable risk factors on cardiovascular disease and mortality. New England Journal of Medicine, 389(14), 1273-1285. URL: https://www.nejm.org/doi/full/10.1056/NEJMoa2206916. DOI: 10.1056/NEJMoa2206916.

Jagannathan, R., Patel, S. A., Ali, M. K., & Narayan, K. V. (2019). Global updates on cardiovascular disease mortality trends and attribution of traditional risk factors. Current diabetes reports, 19, 1-12. URL: https://link.springer.com/article/10.1007/s11892-019-1161-2. DOI: 10.1007/s11892-019-1161-2.

Kallas, D., Sandhu, N., Gandilo, C., Schleicher, M., Banks, L., Jabara, M., ... & Randhawa, V. K. (2023). Use of digital health technology in heart failure and diabetes: a scoping review. Journal of Cardiovascular Translational Research, 16(3), 526-540. URL: https://link.springer.com/article/10.1007/s12265-022-10273-6. DOI: 10.1007/s12265-022-10273-6.

Kedia, S., & Bhushan, M. (2022, June). Prediction of mortality from heart failure using machine learning. In 2022 2nd International Conference on Emerging Frontiers in Electrical and Electronic Technologies (ICEFEET) (pp. 1-6). IEEE. URL: https://ieeexplore.ieee.org/abstract/document/9848348. DOI: 10.1109/ICEFEET51821.2022.9848348.

Kim, Y. G., Choi, Y. Y., Han, K. D., Min, K. J., Choi, H. Y., Shim, J., ... & Kim, Y. H. (2021). Premature ventricular contraction increases the risk of heart failure and ventricular tachyarrhythmia. Scientific reports, 11(1), 12698. URL: https://www.nature.com/articles/s41598-021-92088-0. DOI: 10.1038/s41598-021-92088-0.

Kullo, I. J., & Cooper, L. T. (2010). Early identification of cardiovascular risk using genomics and proteomics. Nature Reviews Cardiology, 7(6), 309-317. URL: https://www.nature.com/articles/nrcardio.2010.53. DOI: 10.1038/nrcardio.2010.53.

Le, M. T., Vo, M. T., Mai, L., & Dao, S. V. (2020, October). Predicting heart failure using deep neural network. In 2020 International Conference on Advanced Technologies for Communications (ATC) (pp. 221-225). IEEE. URL: https://ieeexplore.ieee.org/abstract/document/9255445. DOI: 10.1109/ATC50776.2020.9255445.

Leung, C. K., Fung, D. L., Mushtaq, S. B., Leduchowski, O. T., Bouchard, R. L., Jin, H., ... & Zhang, C. Y. (2020, August). Data science for healthcare predictive analytics. In Proceedings of the 24th Symposium on International Database Engineering & Applications (pp. 1-10). URL: https://dl.acm.org/doi/abs/10.1145/3410566.3410598. DOI: 10.1145/3410566.3410598.

Louridi, N., Douzi, S., & El Ouahidi, B. (2021). Machine learning-based identification of patients with a cardiovascular defect. Journal of Big Data, 8, 1-15. URL: https://link.springer.com/article/10.1186/s40537-021-00524-9. DOI: 10.1186/s40537-021-00524-9.

Lv, J., Dong, B., Lei, H., Shi, G., Wang, H., Zhu, F., ... & Chen, H. (2021). Artificial intelligence-assisted auscultation in detecting congenital heart disease. European Heart Journal-Digital Health, 2(1), 119-124. URL: https://academic.oup.com/ehjdh/article/2/1/119/6066163?login=false. DOI: 10.1093/ehjdh/ztaa017.

Mienye, I. D., Sun, Y., & Wang, Z. (2020). Improved sparse autoencoder based artificial neural network approach for prediction of heart disease. Informatics in Medicine Unlocked, 18, 100307. URL: https://www.sciencedirect.com/science/article/pii/S2352914820300447. DOI: 10.1016/j.imu.2020.100307.

Nalluri, S., Vijaya Saraswathi, R., Ramasubbareddy, S., Govinda, K., & Swetha, E. (2020). Chronic heart disease prediction using data mining techniques. In Data Engineering and Communication Technology: Proceedings of 3rd ICDECT-2K19 (pp. 903-912). Springer Singapore. URL: https://link.springer.com/chapter/10.1007/978-981-15-1097-7_76. DOI: 10.1007/978-981-15-1097-7_76.

Nielsen, D. (2016). Tree boosting with xgboost-why does xgboost win" every" machine learning competition? (Master's thesis, NTNU). URL: https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2433761/16128_FULLTEXT.pdf.

Ramprakash, P., Sarumathi, R., Mowriya, R., & Nithyavishnupriya, S. (2020, February). Heart disease prediction using deep neural network. In 2020 International Conference on Inventive Computation Technologies (ICICT) (pp. 666-670). IEEE. URL:

Rasmy, L., Wu, Y., Wang, N., Geng, X., Zheng, W. J., Wang, F., ... & Zhi, D. (2018). A study of generalizability of recurrent neural network-based predictive models for heart failure onset risk using a large and heterogeneous EHR data set. Journal of biomedical informatics, 84, 11-16. URL: https://www.sciencedirect.com/science/article/pii/S1532046418301175. DOI: 10.1016/j.jbi.2018.06.011.

Reddy, K. S., & Mathur, M. R. (2021). Global Burden of CVD: Prevalence, Pattern, and Trends. In Handbook of Global Health (pp. 423-437). Cham: Springer International Publishing. URL: https://link.springer.com/content/pdf/10.1007/978-3-030-45009-0_24.pdf. DOI: 10.1007/978-3-030-45009-0_24.

Sarma, S., Stoller, D., Hendrix, J., Howden, E., Lawley, J., Livingston, S., ... & Levine, B. D. (2020). Mechanisms of chronotropic incompetence in heart failure with preserved ejection fraction. Circulation: Heart Failure, 13(3), e006331. URL: https://www.ahajournals.org/doi/abs/10.1161/CIRCHEARTFAILURE.119.006331. DOI: 10.1161/CIRCHEARTFAILURE.119.006331.

Srinivas, P., & Katarya, R. (2022). hyOPTXg: OPTUNA hyper-parameter optimization framework for predicting cardiovascular disease using XGBoost. Biomedical Signal Processing and Control, 73, 103456. URL: https://www.sciencedirect.com/science/article/pii/S1746809421010533. DOI: 10.1016/j.bspc.2021.103456.

Valarmathi, R., & Sheela, T. (2021). Heart disease prediction using hyper parameter optimization (HPO) tuning. Biomedical Signal Processing and Control, 70, 103033. URL: https://www.sciencedirect.com/science/article/pii/S1746809421006303. DOI: 10.1016/j.bspc.2021.103033.

Wang, J. (2021, September). Heart failure prediction with machine learning: a comparative study. In Journal of Physics: Conference Series (Vol. 2031, No. 1, p. 012068). IOP Publishing. URL: https://iopscience.iop.org/article/10.1088/1742-6596/2031/1/012068/meta. DOI: 10.1088/1742-6596/2031/1/012068.

Zeng, D., Liu, K., Lai, S., Zhou, G., & Zhao, J. (2014, August). Relation classification via convolutional deep neural network. In Proceedings of COLING 2014, the 25th international conference on computational linguistics: technical papers (pp. 2335-2344). URL: https://aclanthology.org/C14-1220.pdf.