ABUAD Journal of Engineering Research and Development (AJERD)
ISSN (online): 2645-2685; ISSN (print): 2756-6811

Volume 7, Issue 2, 436-445

Early Detection of Congenital Heart Diseases among Infants Using Artificial Neural Network Algorithm

Lucy Ifeyinwa EZIGBO, Anthony KWUBEGHARI, Francis OKOYE

Department of Computer Engineering, Enugu State University of Science and Technology, Enugu State, Nigeria
lucyezigbo752@gmail.com/kwubeghari@gmail.com/francisced@esut.edu.ng

Corresponding Author: lucyezigbo752@gmail.com, +2348032657047
Date Submitted: 23/07/2024
Date Accepted: 02/10/2024
Date Published: 13/10/2024

Abstract: *Congenital Heart Disease (CHD) detection has continued to witness a consistent increase in research attention. CHD diseases are vast and also span across diverse demographics, without sparing pregnant women, unborn babies, or newly born babies. The aim of this study is to develop a detection model capable of detecting heart disease among infants with high accuracy and to also suggest solutions to manage it. To achieve this, the CHD types were identified to develop a data model that considered infants. The data model was processed through imputation, feature selection, and transformation using the imputation method and Principal Component Analysis (PCA). After the data processing stage, Artificial Neural Network (ANN) algorithm is adopted and trained with the data to generate the model for the detection of the CHD. Comparative analysis was used to evaluate the performance of the models in comparison with other models adopted in other works, considering metrics that defined the success of the detection models. The results showed that the ANN has the best detection outcome with an accuracy of 97.44%. Although the use of Logistic Regression algorithm attained a high level of performance with an accuracy of 95.00%, but it still falls below the proposed ANN algorithm. Another highly performing algorithm is the use of Extreme Gradient Boosting (XGB) which achieved an accuracy of 92.00%.*

Keywords*: Congenital Heart Disease (CHD), Infants, Detection Model, Artificial Neural Network, Principal Component Analysis, Extreme Gradient Boosting, Logistic Regression*

1. INTRODUCTION

Congenital heart diseases, more commonly known as Congenital Heart Defects (CHD), represent a diverse and intricate group of structural abnormalities that develop during foetal gestation [1]. These anomalies manifest in various ways and can affect individuals across the entire spectrum of age, from the earliest days of life to adulthood [2]. Notably, the presence of congenital heart diseases can introduce unique challenges, both for those living with the condition and for expectant mothers during pregnancy [3]. CHD is present at birth and results from anomalies in the structure and function of the heart or blood vessels [4]. They are among the most common birth defects, with various forms ranging from minor issues that may not require treatment to complex conditions necessitating surgical intervention [5][6]. The prevalence of CHD varies, but it remains a significant concern globally [7].

According to Wilson et al. [3], CHDs originate during the prenatal phase of development, when the heart is still in its formative stages. These abnormalities occur as the intricate processes that guide the heart's structure and function are disrupted [8][9]. The consequences of CHDs extend across the entire lifespan. Infants born with severe heart defects may require immediate medical attention and surgical intervention, while others with milder conditions might not display symptoms until later in childhood or even into adulthood [10][11]. The degree of severity and the specific type of CHD play a pivotal role in determining the medical journey for each affected individual [12].

CHDs can have a substantial impact on infants and newborns. Some infants may show symptoms immediately after birth, while others may develop issues later in infancy [13]. The severity of the condition plays a critical role in determining the health and development of the child [14]. Timely diagnosis and intervention are essential for ensuring the best outcomes. CHD are of many types, but the three most frequent are Patent Ductus Arteriosus (PDA), Peripartum Cardiomyopathy (PPCM), and Ventricular Septal Defect (VSD).

Artificial Neural Networks (ANNs) are computational models inspired by biological systems, capable of learning relationships between input and output variables through training. ANNs have demonstrated promise in detecting various heart diseases, including Coronary Artery Disease (CAD) and Heart Failure (HF) [15]. For instance, a study conducted by Shah et al. [16] utilized ANN to assess CAD risk by analyzing factors such as age, gender, blood pressure, cholesterol levels, and diabetes.

Various researchers have engaged in the development of an intelligence system aimed at the early detection of hearts diseases, but tackling the issue of CHD in infants have not gotten much attention, hence, this paper focuses on management of Congenital Heart Defects (CHD) affecting infants. This gap when addressed will substantially enrich understanding of the multifaceted landscape of CHD and facilitate the development of an intelligent detective and diagnostic strategy for these intricate cardiac problems.

2. LITERATURE REVIEW

Singh & Kumar [17] proposed a K-Nearest Neighbour (KNN) model for heart disease detection. They utilized the KNN algorithm to classify patient data and achieved an accuracy of 90.79%. The model demonstrated high accuracy but had limitations related to scalability and sensitivity to noisy data. Hossen [18] developed a heart disease prediction model using Support Vector Machine (SVM). Their approach resulted in an accuracy of 83.00%. While SVM provided robust classification performance, it was limited by its dependence on feature scaling and computational complexity. Tiwari et al. [19] implemented a Linear Regression (LR) model for predicting heart disease, achieving an impressive accuracy of 95.00%. Although LR performed well, it was limited by its inability to capture complex relationships in the data compared to more advanced algorithms. However, the study did not consider the early detection of CHD.

Srivenkatesh [20] utilized Extreme Gradient Boosting (XGB) for heart disease prediction, obtaining an accuracy of 91.91%. The XGB model showed high performance and robustness but required extensive tuning of hyperparameters and was computationally intensive. Chithambaram et al. [21] employed Linear Regression (LR) for heart disease prediction with an accuracy of 77.06%. The model was straightforward and easy to interpret but was less effective in handling non-linear relationships and complex data patterns. Chandrika & Madhavi [22] proposed a Gaussian SVM Kernel model, achieving an accuracy of 86.20%. This model offered good performance in capturing non-linear relationships but was computationally demanding and sensitive to parameter tuning. The detection of heart disease among infants was not considered in the study.

Ibrahim et al. [23] developed a Hybrid Random Forest with a Linear Model (HRFLM) for heart disease prediction. Their approach achieved an accuracy of 88.40%. The hybrid model combined the strengths of Random Forest and Linear Regression, though it was complex to implement and required careful calibration of both components. Nandal et al. [24] utilized a combination of KNN and LR models, achieving an accuracy of 86.00%. This approach aimed to leverage the strengths of both models but faced challenges related to integrating the two methods effectively and handling different data characteristics. Pandit [25] employed Extreme Gradient Boosting (XGB) and obtained an accuracy of 92.00%. The model demonstrated strong performance in heart disease prediction but required significant computational resources and tuning to achieve optimal results. However, detection of CHD among infants was not considered in the studies.

3. RESEARCH METHODOLOGY

The methodology used for this research combines experimental and observational methods. The observation method was used for the heart disease data collection, while the experimental method was used for the model implementation, evaluation, and validation. In the realization of the research design, data was collected on CHD, considering classes peculiar to infants. The data was integrated to develop a robust data model for the detection of congenital heart diseases. To ensure the quality of data collection, data processing and analysis were applied to address issues of missing data, data complications, and issues of overfitting that may arise as a result of the utilization of a model for training. In addition, feature selection was applied to allow dimensionality reduction while maintaining data quality, and a feature transformation algorithm was also applied to convert the data into a compatible feature vector for machine learning algorithm identification.

Machine learning algorithm was trained with the data to generate models for the detection of CHD. Comparative analysis was applied to identify the best machine learning-based detection model. A simple prototype model was developed with the model and then tested and validated through practical experimentation.

3.1 Data Collection

The electrocardiogram (ECG) data on congenital heart diseases was collected from four different hospitals in the south-eastern part of Nigeria, considering three major classes: patent ductus artery (PDA), peri partum cardiomyopathy (PPCM), and ventricular septal defect (VSD). The demographics considered for the data collection are preterm infants over a population size of 10 subjects. Data for the PDA class was collected from the Nnamdi Azikiwe University Teaching Hospital, Nnewi, considering a population size of 10 infant subjects, aged 28–38 weeks, using an electrocardiogram machine and a respiratory rate monitor.

3.2 Data Integration

Data integration was a method utilized to merge the multiple datasets collected from the different datasets and data sources and then merge them as one dataset. This was achieved by collecting CHD data from preterm infants, and then heartbeat data from the same demography under normal conditions. In addition, more data was collected from the Physiobank [28] repository and also merged with the existing dataset collected from the primary data sources using the Excel software tool.

3.3 Data Imputation and Normalization

The data collected was first processed through the imputation technique. In this approach, missing data on the dataset was identified and replaced using the mean of the respective variables. This was performed in Excel software and then saved as a CSV file. Having achieved the imputation process successfully, the data normalization approach was applied using the Z-normalization technique. This is a standardization approach that ensures that the data has a mean of zero and a standard deviation of 1. The aim was to ensure that all the features in the dataset are in equal value ranges and also to address issues of classification bias during the training of the algorithm. The model for the standardization was posited in Equation 1 [26];

$$Z_{norm} = \frac{X-mu}{sigma} \quad (1)$$

Where Z is the original data point, mu is the mean of the dataset, sigma is the standard deviation of the dataset.

3.4 Feature Selection

The feature selection algorithm used in this research is the chi-square technique. This approach computes the chi-square between the features of the datasets and identified the number of features with the best chi-square score using the model in Equation 2 [27];

$$X_c^2 = \sum \frac{(O_i - E_i)^2}{E_i} \quad (2)$$

Where c is the degree of freedom, O is the observed feature score, E is the expected feature score and X is the feature selected.

3.5 Feature Transformation

The technique used for the feature transformation approach is the Principal Component Analysis (PCA). To achieve this, the data was standardized with the equation 2; then the Covariance Matrix (C) computed with Equation 3;

$$C = \frac{\Sigma\left((X-\mu) * (X-\mu)'\right)}{(N-1)} \quad (3)$$

Where X is the data matrix where each row represents an observation and each column represents a variable; μ (mu) is the mean vector of the variables; N is the number of data samples. The engine values $E_{vl}$and vectors E_v of the data points are computed with Equation 4 and 5;

$$E_{vl} = det(C - \lambda I) \quad (4)$$

$$\boldsymbol{E_v} = C * v = \lambda * v \quad (5)$$

Where v represents the eigenvector; λ represents the eigen-value associated with that eigenvector; "det" stands for the determinant of the matrix; I is the identity matrix. The PCA algorithm is presented as Algorithm 1; while the PCA life-cycle was reported in Figure 1;

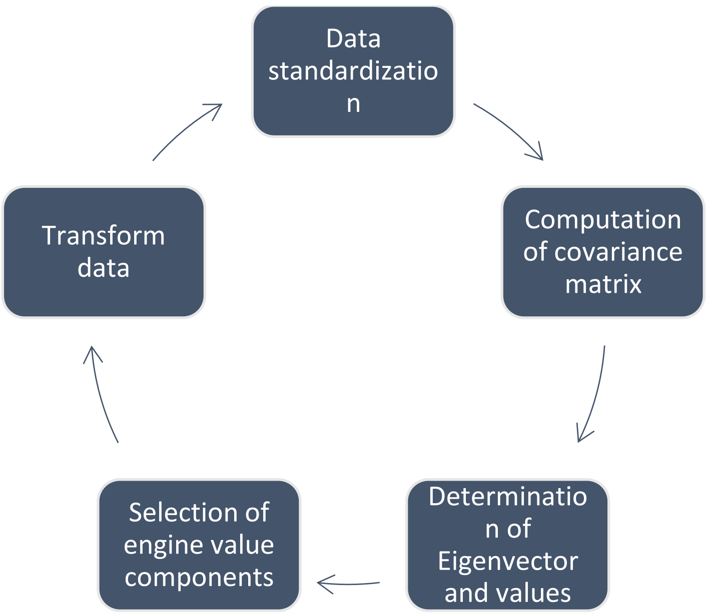


Figure 1: The PCA life-cycle

PCA: Algorithm 1

1. Start
2. Apply equation 2 for data standardize
3. Apply equation 3 for Covariance Matrix computation
4. Calculate Eigen-values and Eigenvectors with equation 4 and 5
5. Sort Eigen-values
6. Select Principal Components
7. Transform Data
8. End

`Figure 2 presents the various steps used in dataset preparation for training. The data after importation was first processed by searching and replacing missing values using the mean imputation technique, then the Z-score standardization technique was applied for the normalization of the data to ensure uniformity in the scaling. Chi-square was then applied for the feature selection of the key attributes with the aim of dimensionality reduction while maintaining data integrity. The data was transformed using the PCA algorithm into a compact feature vector, ready to be applied for the training of machine learning algorithms.

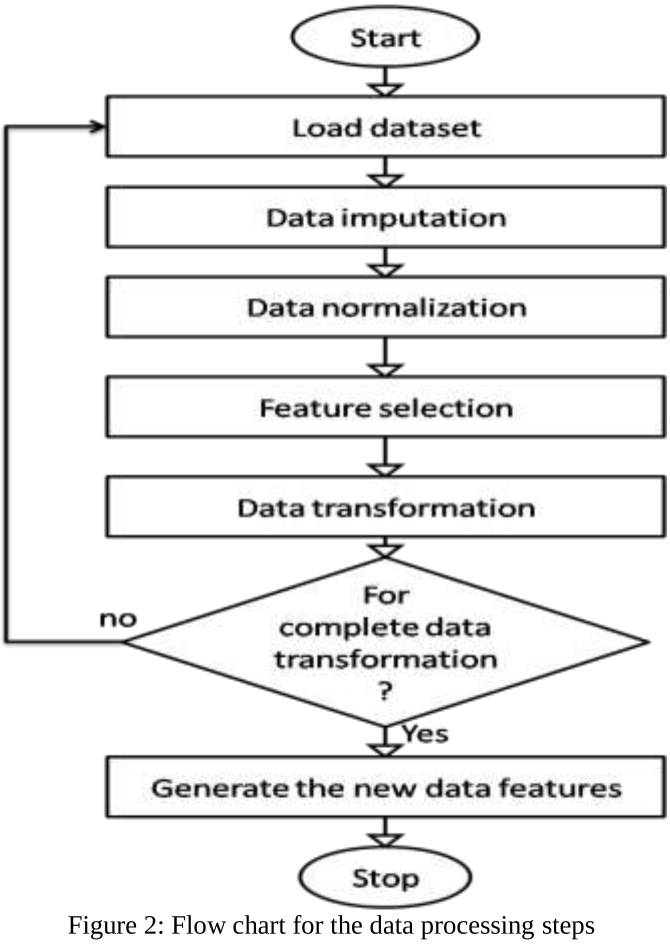


Figure 2: Flow chart for the data processing steps

4. THE ARTIFICIAL NEURAL NETWORK ALGORITHM FOR CHD DETECTION

The Artificial Neural Network (ANN) was utilized in this study as one of the machine learning algorithms selected for the development of the heart disease detection model. The ANN utilized is the feed forward multi layered neural network developed with the interconnected of neurons as depicted in the Figure 3. Where w is the weights, x is the data, b is the bias, Y is the output of the *i* input, where n is the number of inputs to the neurons. The activation function used is the hyperbolic tangent for the normalization of the feature values between 0 and 1. The interconnected neural network was presented in the Figure 4. The Figure 4 presented the architectural model of the neural network with three hidden layers, input layers and output layers for the detection. The input layers were constituted of the features of 20 features which made of the CHD data collected considering the demography (infants). The output on the other hand is a binary classification which detects normal or abnormal CHD.

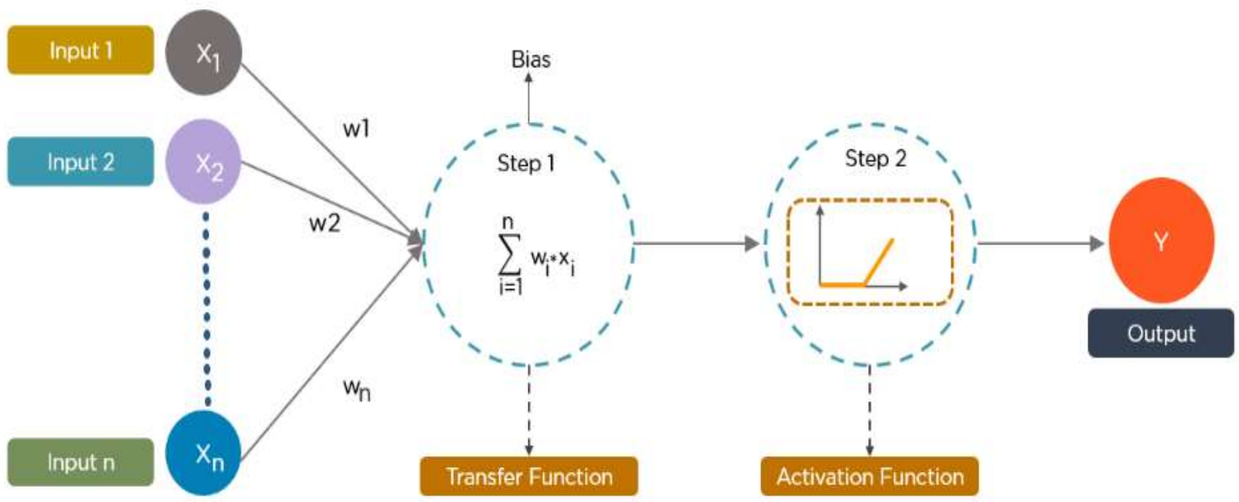


Figure 3: Neural network architecture

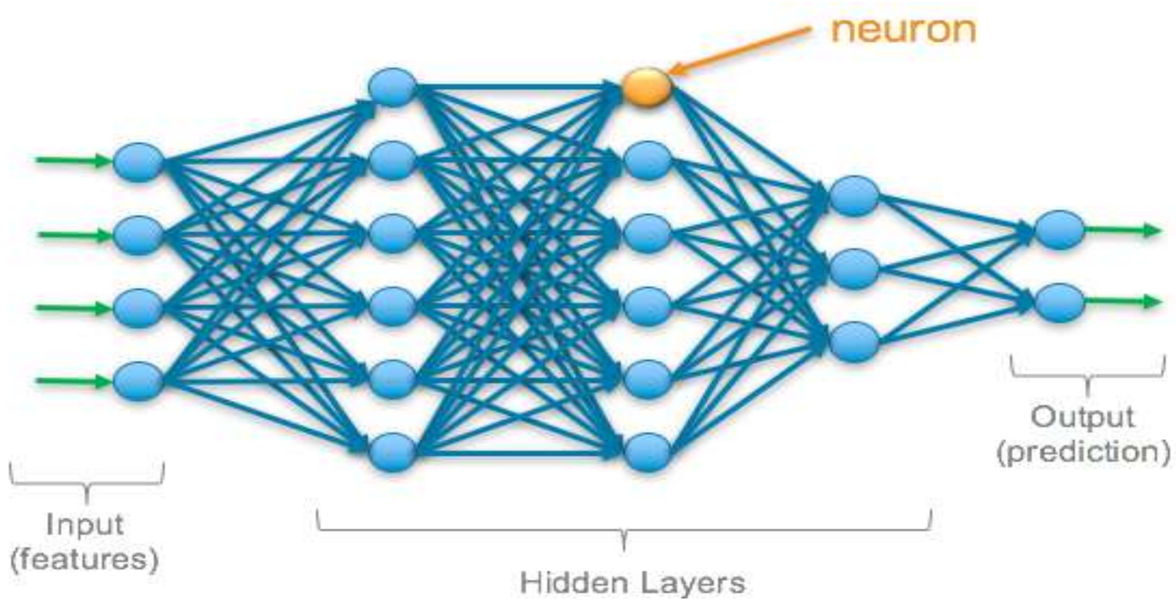


Figure 4: Architecture of the feed-forward multi layered neural network

Algorithm 2: The Neural Network Algorithm

1. Start
2. Initialization are hyper-parameters
3. Define the network architecture
4. Receive input data as a feature vector.
5. For each neuron
6. Passes the input values forward
7. Compute the weighted sum of inputs for each neuron
8. Weighted Sum = $b + \sum_{i=1}^{n} w_i x_i$
9. Apply an activation function to the weighted sum as $(b + Tanh\ \sum_{i=1}^{n} w_i x_i)$
10. Pass the outputs from one layer as inputs to the next layer.
11. Repeat this process for each layer until the output layer is reached.
12. Get final Output
13. End

5. DEVELOPMENT OF THE HEART DISEASE DETECTION (HDD) MODEL USING ANN TECHNIQUE

To achieve this, artificial neural network was trained respectively to generate the heart disease detection model. The machine learning algorithm was independently trained to create a robust and accurate heart disease detection model. By leveraging the strengths of the algorithm, aimed to enhance our detective capability, ultimately contributing to improved heart disease diagnosis and risk assessment.

5.1 Training of the ANN Algorithm

The training of the ANN algorithm was achieved utilizing the dataset which was divided into training, test and validation sets in the ratio of 80:10:10 and then trained with back-propagation optimization [27] process, which adjusts the neurons hyper-parameters, while monitoring the loss function until the detection model is generated. The training process was achieved as shown in the flow chart of Figure 5 to generate the detection model for heart disease.

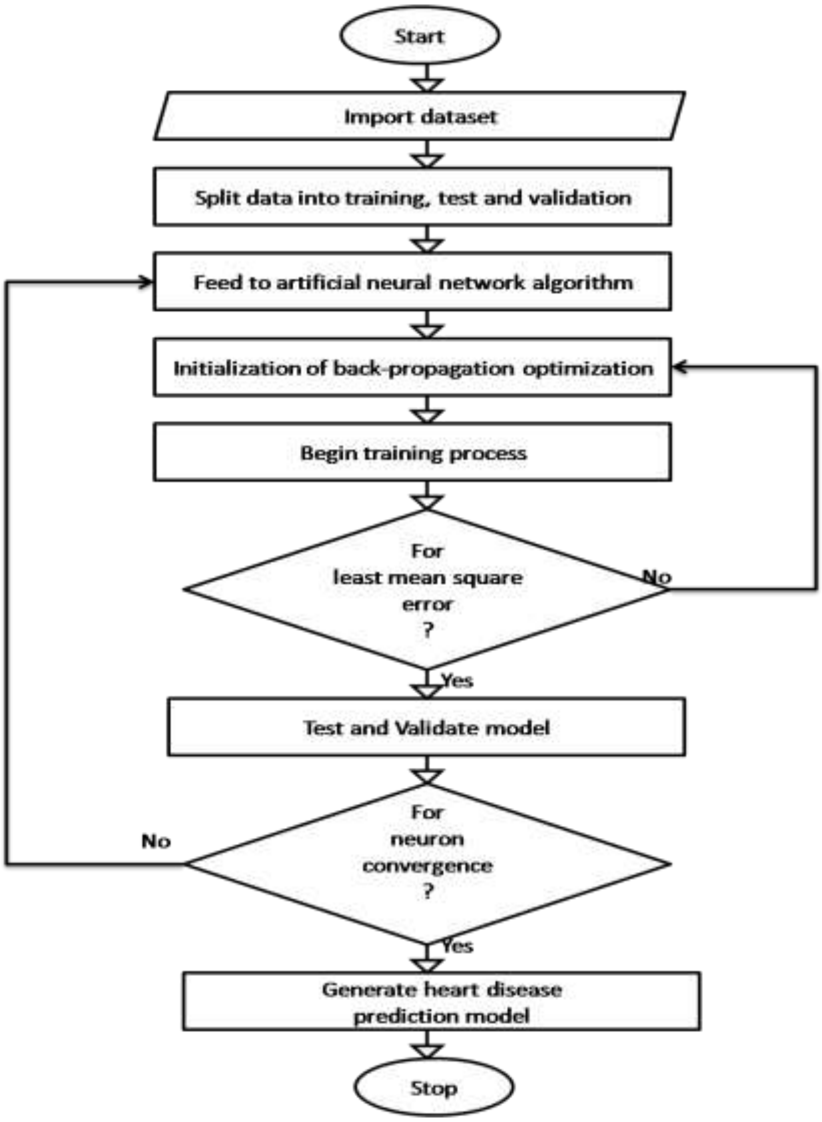


Figure 5: Flowchart of the ANN training process

Figure 5 presents the flow chart of the trained ANN algorithm for the generation of the HDD. The processed data using imputation, the Z-normalization model in equation 1, and the PCA-based feature selection algorithm was imported into the neural network algorithm and then trained using the back-propagation process, which considers the gradient loss of the neurons and then adjusts the neuron hyper-parameters during the training process until neurons converge after testing and validation before generating the heart disease detection model.

4.2 The Machine Learning Based Heart Disease Detection Model

The machine learning-based heart disease detection model was developed using a trained algorithm and a test dataset of normal and abnormal heart features. The data was imported into the data processing tool for feature selection, considering key features with the highest-ranking score as depicted in the chi-square feature selection model of Equation 2. The PCA algorithm was then applied to convert these features into a compact feature vector for compatibility with the trained machine learning algorithms for heart disease detection. The test features are identified and then compared with the trained algorithm features, whose outcomes are classified under the class of heart diseases and return heart disease detection; however, when the features are classified as normal heart beat, the outcome is returned as normal heart beat. The flowchart in Figure 6 was utilized in presenting the data flow of the HDD.

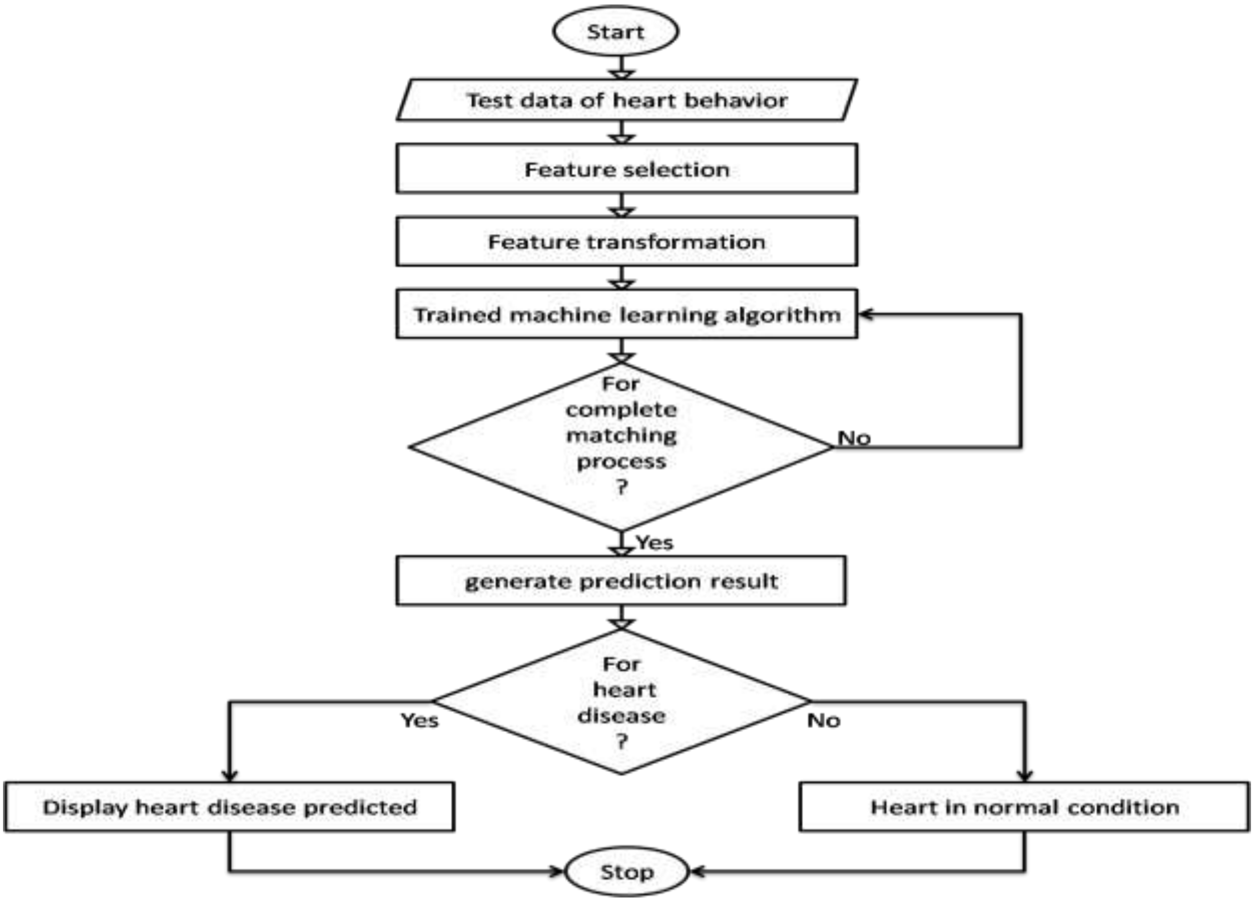


Figure 6: Flowchart of the heart disease detection model

Figure 6 shows how the test data were imported for processing using the feature selection approach in the chi-square model of Equation 2. The selected features are the key attributes that model the main behaviour of heart disease instances. These data are transformed using the PCA algorithm and then fed to machine learning for training and the generation of the detection result. When the detected result is heart disease, it returns an output to inform the examiner of the detection outcome, while the outcome when heart disease is not detected is also returned as a normal heartbeat condition.

6. SYSTEM RESULTS

This section presents the results of the ML training process, considering the RF, DT, and ANN trained with the heart detection model. The results consider performance indicators for detection models such as residual distribution, residual plots, regression, mean absolute error, root mean square error, and mean square error to analyze the models, respectively. This plot result considered in this study was used to measure the pattern of actual data distribution across each model during the training process, and the aim was to ensure that the errors were normally distributed and also to see if there was a pattern in the distribution of actual data. This behaviour of the residual across the model helps deduce the strength or issues in the capacity of the model to perform detection. In the ANN validation, Table 1 was applied.

Table 1: Validation of the ANN model for heart disease detection

Iteration	MSE	RMSE	MAE	Accuracy
1	0.08514	0.2861	0.2436	0.9744
2	0.08564	0.2774	0.2563	0.9863
3	0.08763	0.2646	0.2640	0.9845
4	0.08654	0.2345	0.2535	0.9844
5	0.08532	0.2447	0.2644	0.9747
6	0.08323	0.2534	0.2535	0.9893
7	0.08644	0.2643	0.2645	0.9845
8	0.08567	0.2622	0.2533	0.9917
9	0.08646	0.2645	0.2567	0.9860
10	0.08656	0.2543	0.2035	0.9744
Average	0.08586	0.2606	0.25133	0.9895

Table 1 presents the results of the ANN validation considering the MSE, RMSE, and MAE. From the results, it was observed that the average MSE is 0.085856, the RMSE is 0.2606, and the and the MAE is 0.25133 and 0.9895, respectively. These results implied that the ANN was able to achieve a limited error between the actual and detected values. Overall, the performance of the artificial ANN in detecting the values was satisfactory, with relatively low MSE, RMSE, and MAE values. This indicates that the ANN model effectively captured the underlying patterns in the heart disease data and produced detection that were close to the actual values.

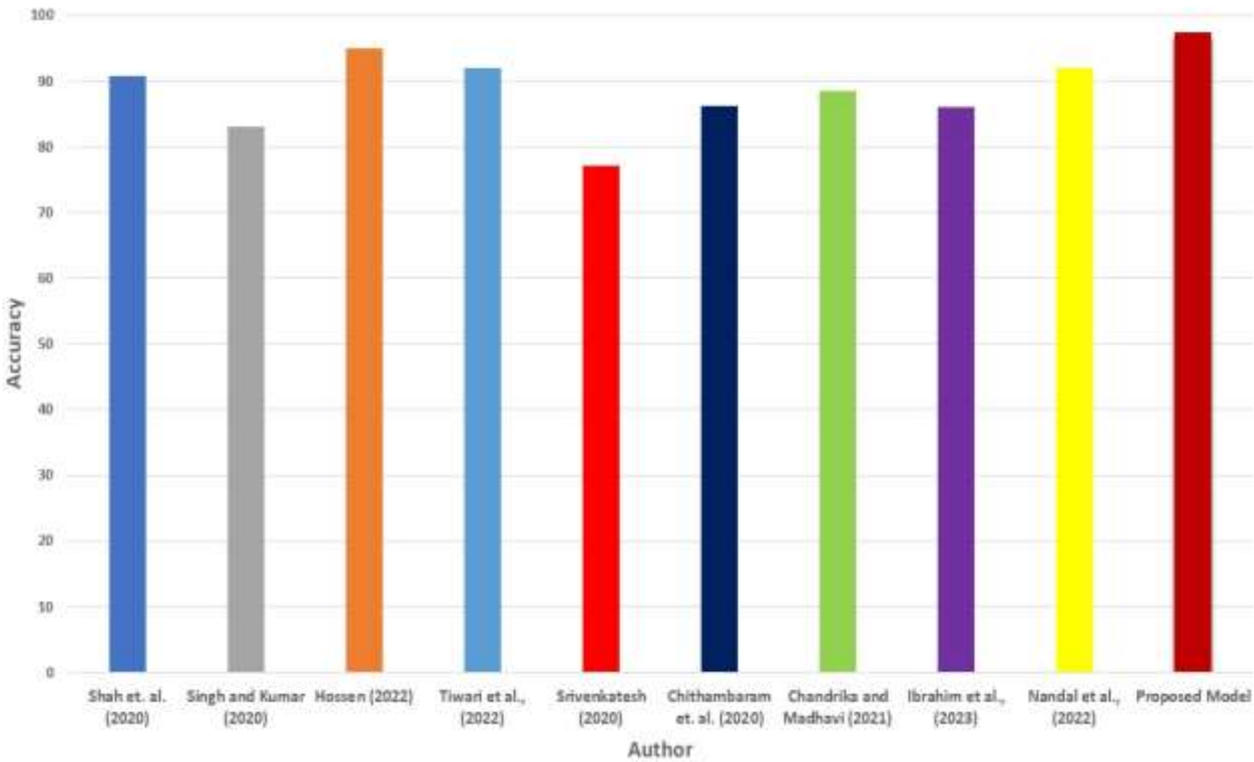
6.1 Comparative Analysis

The comparative analysis considered the various models and the accuracy result of the models. The results of the comparison are presented after cross-validation and reported in Table 2.

Table 2: Comparative analysis of machine learning algorithms for heart disease detection

Authors	Machine Learning Algorithm	Performance Accuracy (%)
Singh & Kumar [17]	K-Nearest Neighbour (KNN)	90.79
Hossen [18]	Support Vector Machine	83.00
Tiwari et al. [19]	Linear Regression (LR)	95.00
Srivenkatesh [20]	Extreme Gradient Boosting (XGB)	91.91
Chithambaram et al. [21]	LR	77.06
Chandrika & Madhavi [22]	Gaussian SVM Kernel	86.20
Ibrahim et al. [23]	Hybrid Random Forest with a Linear Model (HRFLM)	88.40
Nandal et al. [24]	KNN & LR	86.00
Pandit [25]	XGB	92.00
Proposed Model	ANN	97.44

The comparative results presented are being illustrated in Figure 7 for a better representation of the outcome of the comparative analysis.



From the comparative analysis presented in Figure 7, it can be observed that out of all the models presented, the ANN algorithm turns out to be the highest performing algorithm with an accuracy of 97.44% when compared to the other algorithms. Although the use of Logistic Regression algorithm by [19] attained a high level of performance with an accuracy of 95.00%, but it still falls below the proposed ANN algorithm. Another highly performing algorithm is the use of XGB by [20] and [25] which achieved an accuracy of 92.00% and 91.91% respectively. However, these algorithms show significant accuracy in terms of performance in detecting heart diseases, but the use of ANN has proven to be the most efficient and is recommended for future application in heart disease diagnostic systems. Generally, while it is possible that other literatures may superseded this accuracy which the researcher is not aware of, the selling point of the model is the demography considered for data collection, which is lacking in previous literatures. However, the model is may not be applicable for the management of other types of heart diseases, other than CHD, and this is the limitation of the study.

7. CONCLUSION

Over the years, congenital heart disease (CHD) detection has continued to witness a consistent increase in research attention. CHD diseases are vast and also span across diverse demographics, without sparing pregnant women, unborn babies, or newly born babies. While several studies have been presented considering other demography, this three aforementioned demography have received very little attention despite the high mortality rate and rising issues of heart disease among the categories of persons. The aim of this study is to develop a detection model capable of detecting heart disease among infants with high accuracy and to also suggest solutions to manage it. To achieve this, the CHD was characterized and the types were identified to develop a data model that considered infants. The data model was processed through imputation, feature selection, and transformation using the imputation method and PCA. After the data processing stage, Artificial Neural Network (ANN) algorithm is adopted and trained with the data to generate the model for the detection of the CHD.

Comparative analysis was used to evaluate the performance of the models in comparison with other models adopted in other works, considering metrics that defined the success of the detection models. The results showed that the ANN has the best detection outcome with an accuracy of 97.44%. Although the use of Logistic Regression algorithm by [19] attained a high level of performance with an accuracy of 95.00%, but it still falls below the proposed ANN algorithm. Another highly performing algorithm is the use of XGB by [20] and [25] which achieved an accuracy of 92.00% and 91.91% respectively.

REFERENCES

[1] Zhu, X., Wang, X., Zhang, S., & Zhang, Y. (2019). Heart disease prediction using machine learning techniques. *Journal Informatics, 98*, 103-115. https://doi.org/10.1016/j.jbi.2019.103115

[2] Yang, X., Li, Y., Li, Y., Ren, X., Zhang, X., Hu, D., & Sun, Y. (2017). Predictive models for the risk of coronary artery disease among Chinese people. *Scientific Reports, 7*(1), 1-9. https://doi.org/10.1038/s41598-017-15138-3

[3] Wilson, P. W., D'Agostino, R. B., Levy, D., Belanger, A. M., Silbershatz, H., & Kannel, W. B. (1998). Prediction of coronary heart disease using risk factor categories. *Circulation, 97*(18), 1837-1847. https://doi.org/10.1161/01.CIR.97.18.1837

[4] World Health Organization (WHO). (2023). A woman dies every two minutes due to pregnancy or childbirth. Geneva/New York/Washington. Retrieved from https://www.who.int/news/item/23-02-2023-a-woman-dies-every-two-minutes-due-to-pregnancy-or-childbirth--un-agencies

[5] Singh, N., & Jindal, S. (2018). Heart disease prediction system using hybrid technique of data mining algorithms. *International Journal of Advance Research, Ideas and Innovations in Technology, 4*(2), 15-18. https://www.ijarai.thesai.org

[6] Adil, H. S., & Chaurasia, P. K. (2019). A review on heart disease prediction using machine learning techniques. *International Journal of Management, IT & Engineering, 9*(4), 208-224.

[7] Juma, P. A., Mohamed, S. F., Matanje-Mwagomba, B. L., et al. (2018). Non-communicable disease prevention policy process in five African countries. *BMC Public Health, 18*(Suppl 1), 961. https://doi.org/10.1186/s12889-018-5830-6

[8] Shirwalkar, N., & Tak, T. (2018). Human heart disease prediction system using data mining techniques. *International Journal of Innovations & Advancement in Computer Science, 7*(3), 101-108.

[9] Nnenna, H. N., Odo, H., Ozoemena, P., & Ebere Uzoka Chidi. (2022). Predictive model for the monitoring and detection of heart disease using wavelet-based machine learning technique. *American Journal of Applied Sciences and Engineering*, 3(5), 1-12. https://doi.org/10.21203/rs.3.rs-97004/v1

[10] Arabasadi, Z., Alizadehsani, R., Roshanzamir, M., Moosaei, H., & Yarifard, A. A. (2017). Computer aided decision making for heart disease detection using hybrid neural network-genetic algorithm. *Computer Methods and Programs in Biomedicine, 141*, 19-26. https://doi.org/10.1016/j.cmpb.2017.01.007

[11] Kohil, P. (2023). Everything you need to know about heart disease. Available at: https://www.medicalnewstoday.com/articles/237191. Retrieved on September 28, 2023.

[12] Adeniyi, K. M., Oladapo, O. J., Araoye, T. O., Adebayo, T. F., Egoigwe, S. V., & Odo, M. C. (2023). Prediction of heart disease outcomes using machine learning classifier. *Indonesian Journal of Electrical Engineering and Computer Science, 30*(2), 917-926. https://doi.org/10.11591/ijeecs.v30.i2.33207

[13] Manrique, A., Carrasco, S., & Buil, P. (2019). Machine learning in cardiovascular medicine: Are we there yet? *Revista Española de Cardiología (English Edition), 72*(1), 81-83. https://doi.org/10.1016/j.rec.2018.07.013

[14] Patel, J., Upadhyay, T., & Patel, S. (2018). Heart disease prediction using machine learning and data mining technique. *IJCSC, 7*(1), 129-137. https://www.ijcsc.org

[15] Roger, V. L. (2021). Epidemiology of heart failure: A contemporary perspective. *Circulation Research, 128*(10), 1421-1434. https://doi.org/10.1161/CIRCRESAHA.121.318365

[16] Shah, D., Patel, S., & Bharti, K. S. (2020). Heart disease prediction using machine learning techniques. *SN Computer Science, 1*, 345. https://doi.org/10.1007/s42979-020-00365-y

[17] Singh, A., & Kumar, R. (2020). Heart disease prediction using machine learning algorithms. In *2020 International Conference on Electrical and Electronics Engineering (ICE3-2020)*. https://doi.org/10.1109/ice3.2020.9280795
[18] Hossen, M. (2022). Linear regression-based heart disease prediction using machine learning. *Journal of Health Informatics and Statistics, 7*(1), 55-68. https://doi.org/10.1007/s41899-022-00109-0
[19] Tiwari, A., Chugh, A., & Sharma, A. (2022). Ensemble framework for cardiovascular disease prediction. In *2022 International Conference on Machine Learning and Data Science (MLDS)*. https://doi.org/10.1109/mlds54165.2022.9882904
[20] Srivenkatesh, M. (2020). Prediction of cardiovascular disease using machine learning algorithms. *International Journal of Engineering and Advanced Technology (IJEAT), 9*(3), 2249-8958. https://doi.org/10.35940/ijeat.C5864.059320
[21] Chithambaram, T., Kannan, N. L., & Gowsalya, M. (2020). Heart disease detection using machine learning. *ResearchSquare*. https://doi.org/10.21203/rs.3.rs-97004/v1
[22] Chandrika, L., & Madhavi, K. (2020). A hybrid framework for heart disease prediction using machine learning algorithms. *E3S Web of Conferences, 309*, 01043. https://doi.org/10.1051/e3sconf/202130901043
[23] Ibrahim, S., Salhab, N., & Falou, A. E. (2023). Heart disease prediction using machine learning. In *2023 1st International Conference on Advanced Innovations in Smart Cities (ICAISC)*. https://doi.org/10.1109/icaisc56366.2023.10085522
[24] Nandal, N., Goel, L., & Tanwar, R. (2022). Machine learning-based heart attack prediction: Asymptomatic heart attack prediction method and exploratory analysis. *F1000Research, 11*, 1126. https://doi.org/10.12688/f1000research.29083.1
[25] Pandit, B. (2023). Four most popular data normalization techniques every data scientist should know. Available at https://dataaspirant.com/data-normalization-techniques/. Accessed October 23, 2023.
[26] Al-Garad, M. (2020). Machine learning techniques for heart disease prediction. *Journal of Data Science and Analytics, 4*(2), 99-108. https://doi.org/10.1007/s10115-020-04896-7
[27] Reyna, M., Kiarashi, Y., Elola, A., Oliveira, J., Renna, F., Gu, A., Perez Alday, E. A., Sadr, N., Mattos, S., Coimbra, M., Sameni, R., Bahrami Rad, A., Koscova, Z., & Clifford, G. (2023). Heart murmur detection from phonocardiogram recordings: The George B. Moody PhysioNet Challenge 2022 (version 1.0.0). *PhysioNet*. https://doi.org/10.13026/t49p-5v35
[28] PhysioNet. (2019). PhysioBank databases. Retrieved from https://physionet.org Accessed in September 2024