Comprehensive Comparative Analysis of Computational Intelligence Models for Heart Disease Prediction

Kange Theophilus Aondofa(1)*
Abdullahi Isa(1)

Received: 13/10/2024
Revised: 11/11/2024
Accepted: 17/12/2024

---

[1] Department of Computer Science, Faculty of Physical Sciences, University of Maiduguri, Nigeria
* Corresponding Author's Email: isaabdullahi@unimaid.edu.ng

https://doi.org/10.20428/jst.v30i1.2512

Comprehensive Comparative Analysis of Computational Intelligence Models for Heart Disease Prediction

Kange Theophilus Aondofa
Department of Computer Science, Faculty of Physical Sciences, University of Maiduguri, Nigeria
jandandewizard@gmail.com

Abdullahi Isa
Department of Computer Science, Faculty of Physical Sciences, University of Maiduguri, Nigeria
isaabdullahi@unimaid.edu.ng

Abstract— Heart disease is a major global public health concern, highlighting the urgent need for precise prediction models to enhance prevention and early detection. This research looks at how well different computer models can predict heart disease. It mainly looks at Random Forests (RF), Logistic Regression (LR), Convolutional Neural Networks (CNNs), Decision Trees (DTs), and Artificial Neural Networks (ANNs). According to our data, ANNs were the most accurate model for prediction, achieving the highest accuracy (87.0%) with balanced recall (0.84), strong precision (0.91), and F1-score (0.87). CNNs displayed solid overall performance with an accuracy of 84.0%, precision and recall both at 0.85, and an F1 score of 0.85, suggesting their effectiveness in learning spatial hierarchies. Despite logistic regression attaining an accuracy of 82.0%, its precision (0.89) and recall (0.79) indicate a compromise between accurately detecting genuine positives and reducing false positives. Both Decision Trees and Random Forests achieved perfect precision, recall, and F1-scores (all 1.00), though Decision Trees had a lower accuracy of 74.0%, potentially indicating overfitting. This comparative analysis highlights the advantages and disadvantages of each model, offering insights into their usefulness for predicting heart disease. By incorporating essential metrics such as accuracy, precision, recall, and F1-score, this research contributes to creating more precise and dependable diagnostic tools, setting the stage for improved prevention and early intervention tactics for heart disease.

Keywords— Artificial neural networks, heart disease prediction, Machine learning algorithms, Analysis of AI Models, Computational intelligent models, Deep learning Models.

I. INTRODUCTION

Cardiovascular diseases (CVD), especially heart disease, continue to be a prominent cause of mortality around the world. The World Health Organization (WHO) reports that heart-related issues result in nearly 17.9 million deaths annually, with over 80% of these fatalities stemming from heart attacks and strokes [1], [2]. Cardiovascular disease (CVD) encompasses various disorders, such as coronary heart disease, cerebrovascular disease, peripheral arterial disease, rheumatic heart disease, and congenital heart disease. Unhealthy choices in lifestyle, like poor nutrition, lack of physical activity, excessive alcohol consumption, and smoking, elevate the risk of heart disease by contributing to high blood pressure, increased blood glucose levels, high cholesterol, and obesity [3]. Early detection of individuals at high risk for CVD and timely action can help avert premature deaths. In resource-constrained environments, particularly in low- and middle-income countries (LMICs), financial constraints and limited access to healthcare services pose significant challenges to delivering affordable and accurate diagnoses [4]. The global economic burden of cardiovascular diseases is immense, with an estimated cost of USD 3.7 trillion between 2010 and 2015 [3].

Despite the increasing burden of CVD, there is a growing need for efficient, cost-effective, and accurate diagnostic tools to aid in early diagnosis and prevention, particularly in regions with constrained healthcare infrastructure. Misclassifying heart disease in patients as negative has more severe consequences than false positives, necessitating reliable prediction systems [5]. Recent years have seen a rise in the use of machine learning (ML) [6, [7], [8], [9], [10], [11], [12] and deep learning (DL) models [11], [13], [14], [15], [16], [17] to predict heart disease. These models have the potential to revolutionize healthcare by providing precise, efficient, and scalable diagnostic solutions.

Traditionally, heart disease prediction has depended on clinical assessments and diagnostic tools like electrocardiograms (ECGs). However, manually evaluating long-term ECG data takes a lot of effort and is prone to mistakes. Researchers have created machine learning-based heart disease detection (MLBHDD) systems as adaptable and affordable substitutes in light of the growing use of ML in healthcare [18]. In order to evaluate patient data and identify those who are at risk of heart disease, these systems use predictive algorithms. DL models, in particular, have demonstrated virtually 100% accuracy on some datasets, which is a promising development for the diagnosis of heart disease [19].

However, there are still major obstacles to overcome. Lack of interpretability is a major problem with ML and DL techniques. Neural networks with several hidden layers are examples of complex DL models that frequently function as "black boxes," making it challenging to comprehend how they make decisions. This opacity raises questions about these models' reliability and accountability in therapeutic settings where explainability and openness are crucial [19]. Additionally, many ML models suffer from bias toward the majority class in imbalanced datasets, resulting in poor performance when predicting outcomes for minority classes. This imbalance is particularly problematic in medical data, where certain conditions or outcomes may be underrepresented [3]. Consequently, there is a need to address these issues and develop more interpretable, fair, and unbiased heart disease prediction models.

Numerous studies have looked at various machine learning and deep learning techniques in the application of computational intelligence models in the prediction of heart disease. In order to predict cardiac disease using clinical data, early research in this area concentrated on conventional machine learning methods like logistic regression (LR) and decision trees (DT). These models offer simplicity and interpretability, making them easy to deploy in healthcare settings. However, their predictive performance is often limited by the complexity of heart disease, which involves numerous interacting risk factors [3].

Advanced machine learning models like Random Forests (RF) and Artificial Neural Networks (ANNs), which are skilled at identifying non-linear patterns in data, have become the focus of recent research.

Random Forests, an ensemble technique, have shown particular effectiveness in minimizing overfitting and enhancing generalization. Meanwhile, ANNs, modeled after the human brain, excel at learning complex patterns in large datasets [18]. Although ANNs are frequently criticized for their lack of interpretability and demand significant processing resources, studies have demonstrated that they can reach excellent accuracy in the prediction of cardiac disease [19].

The discipline has seen significant transformation with the introduction of deep learning techniques, especially Convolutional Neural Networks (CNNs). CNNs were first created for image processing, but they have now been modified for time-series research, including the interpretation of ECG signals, and have demonstrated remarkable predictive power [5]. Despite their success, CNNs share the interpretability issues of other deep learning models, as their decision-making processes are difficult to explain. Moreover, while DL models have achieved near-perfect accuracy in some cases, their performance can be inconsistent across different datasets, raising concerns about their generalizability [18].

Numerous studies have investigated the effect of data imbalance on ML model performance. Research by [3] highlighted the challenges of imbalanced datasets, were dominant majority classes often skew predictions, resulting in poor recall for minority classes. Techniques including oversampling, undersampling, and synthetic data synthesis (like SMOTE) have been investigated to lessen this, but the results have been inconsistent. While these methods can enhance model performance, they may also introduce new issues, such as overfitting or biased predictions [4]. Despite these developments, there are still a lot of unanswered questions in the field. Very few papers offer thorough analyses of several ML and DL models designed especially for the prediction of heart disease, especially when it comes to comparing performance across crucial metrics like accuracy, precision, recall, and F1-score. Additionally, little study has been done on how fair and interpretable these models are in actual clinical settings. This work aims to close these gaps by evaluating the advantages and disadvantages of several computational intelligence models in the prediction of cardiac disease.

Given the growing prevalence of heart disease and the increasing reliance on computational models for its prediction, there is a pressing need for a systematic assessment of these models. While numerous studies have explored individual ML and DL models, few have provided a comprehensive comparison of their performance across multiple metrics. Moreover, issues such as interpretability, fairness, and bias remain underexplored. By evaluating and contrasting the efficacy of several computational intelligence models—including Logistic Regression, Convolutional Neural Networks Random Forests, Artificial Neural Networks, and Decision Trees—in the prediction of heart disease, this study seeks to address these issues.

The main objective is to examine these machine learning (ML) and deep learning (DL) models in detail and compare their performance using metrics like accuracy, precision, recall, and F1-score. The study will investigate each model's interpretability, fairness, and potential biases in addition to performance indicators, especially when dealing with imbalanced datasets. This study aims to provide important insights into each approach's practical usefulness by highlighting its advantages and disadvantages. The study will also highlight the difficulties and possibilities for improving the precision and efficacy of computational intelligence models in the prediction of cardiac disease.

II. METHODOLOGY

Transparency and reproducibility are ensured by the thorough explanation of the study's methods in this section. From data collection to model evaluation, the study was conducted in a methodical manner. Every stage was thoughtfully planned to improve the validity and dependability of the outcomes. The variables used in the study are defined conceptually and operationally to help readers understand the scope and relevance of the research. This methodology includes information about the dataset, data processing, feature selection, model evaluation, and validation techniques.

A. Dataset Source

The Heart Disease Dataset, which was used in this study, was obtained from the Kaggle platform [18]. Although only the Cleveland database was utilized for this experiment, this dataset combines four distinct datasets. It is an open-access dataset that includes several heart disease-related characteristics that, according to other studies, are important predictors [20], [21]. The dataset contains medical records for 303 patients, each with 14 clinical attributes that were identified as relevant for predicting heart disease.

B. Dataset Description

The dataset used contains 14 features, explained in Table 1. These features provide comprehensive clinical data related to patient health, ranging from demographic information to results from various medical tests. The **Target** variable is binary, indicating whether a patient has heart disease (1) or not (0)

C. Data Processing

Data preprocessing was conducted to clean and transform the dataset for machine learning. Multiple steps were undertaken, as outlined below:

1) *Data Augmentation:* The dataset, initially containing 303 records, was deemed insufficient for robust machine learning models. As a result, random number generation techniques were employed, utilizing the minimum and maximum values of each attribute to enhance the dataset. This step tripled the dataset's size, improving model training and reducing bias [3]. This augmentation positively impacted the classifiers' performance, as observed in the results section.
2) *Data Type Transformation:* Some attributes were transformed to fit model requirements. Since this study used binary classification, certain attributes were converted into categorical or numerical formats where appropriate. The Target attribute, used for binary classification, was 0 (no heart disease) and 1 (heart disease). Independent variables like chest pain type and thalassemia were encoded to align with the machine learning models' specifications.
3) *Nominal Attributes:* Certain features such as Thal, CP, and Major Vessels were nominal, and their values were encoded as integers. For example, the Thal attribute had predefined values representing different conditions, and similar encoding was applied to other nominal features.

TABLE 1. DATASET INFORMATION

Attribute Name	Description
Age	Age of the patient (years)
Sex	Patient's sex (1 = male, 0 = female)
Chest Pain Type (CP)	Type of chest pain (1: typical angina, 2: atypical angina, 3: non-anginal, 4: asymptomatic)
Resting Blood Pressure	Blood pressure at rest (mm Hg)
Serum Cholesterol	Cholesterol level (mg/dL)
Fasting Blood Sugar	Fasting blood sugar (>120mg/dL, 1 = true; 0 = false)
Resting ECG Results	ECG results (0 = normal, 1 = ST-T wave abnormality, 2 = probable/definite left ventricular hypertrophy)
Max Heart Rate	Maximum heart rate achieved
Exercise Induced Angina	Presence of exercise-induced angina (1 = yes, 0 = no)
Oldpeak	ST depression relative to rest
ST Segment	Slope of the ST segment during peak exercise (1 = upsloping, 2 = flat, 3 = downsloping)
Major Vessels	Number of major vessels (0-3)
Thal	Thalassemia (3 = normal, 6 = fixed defect, 7 = reversible defect)
Heart Disease [Target]	Whether the patient has heart disease (1 = yes, 0 = no)

D. Data Cleaning

Data cleaning was essential to remove noise and inconsistencies. The cleaning process involved:

- Eliminating duplicates to avoid skewing the results.
- Handling missing values by using appropriate imputation techniques.

- Removing irrelevant information that did not contribute to the prediction of heart disease.

1) Data Augmentation: Feature selection plays a vital role in enhancing model accuracy by removing irrelevant or redundant features [22]. In this investigation:

a) The most relevant features were chosen based on their significance in predicting heart disease.
b) Irrelevant features were discarded to reduce dimensionality and improve computational efficiency.
c) This step enhanced model performance by simplifying the data while retaining essential information.

The feature selection process helped:

- Reduce model complexity, thereby speeding up the training process.
- Avoid overfitting by eliminating noise from the dataset.

E. Model Evaluation and K-Fold Cross-Validation

The machine learning models were evaluated using **K-fold cross-validation** to avoid bias and ensure generalizability. A 10-fold cross-validation technique was employed (Alotaibi, 2019), in which the dataset was randomly partitioned into 10 equal-sized subsets. Each subset was used once as a test set, while the remaining subsets served as the training set. By reducing overfitting and underfitting, this method produced a reliable evaluation of the model's performance.

The dataset was divided into 70% training data and 30% testing data in order to further validate the findings. By ensuring that models were exposed to unseen data during evaluation, this divide provided insights into the models' practicality.

III. RESULTS AND DISCUSSION

This section summarizes our research's results and offers a thorough analysis of them. The accuracy, precision, recall, and F1-score of five distinct machine learning models—Artificial Neural Networks (ANNs), Convolutional Neural Networks (CNNs), Logistic Regression (LR), Decision Trees (DT), and Random Forests (RF)—were assessed. To evaluate each model's classification performance on the heart disease dataset, several metrics were calculated.

TABLE 2. MODEL PERFORMANCE COMPARISON

Model	Accuracy	Precision	Recall	F1-score
ANNs	87.0%	0.91	0.84	0.87
CNNs	84.0%	0.85	0.85	0.85
LR	82.0%	0.89	0.79	0.84
DT	74.0%	1.00	1.00	1.00
RF	82.1%	1.00	1.00	1.00

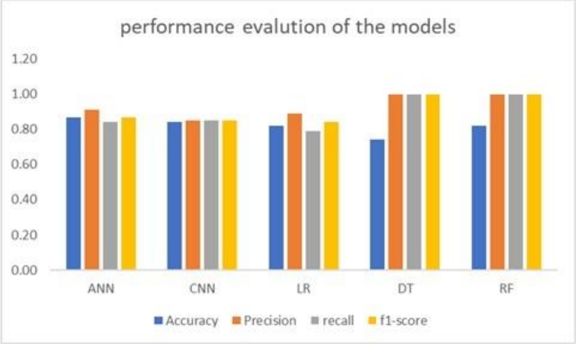

Fig. 1. Model Performance Evaluation Metrics

A. Model Performance Analysis

1) Accuracy: Accuracy measures the overall correctness of the models by calculating the ratio of correctly predicted instances to the total instances. As shown in Table 2, ANNs achieved the highest accuracy (87.0%), followed closely by CNNs (84.0%) and Random Forests (82.1%). Decision Trees (74.0%) performed the worst in terms of accuracy, indicating that simple decision trees might be overfitting or failing to generalize well to unseen data.

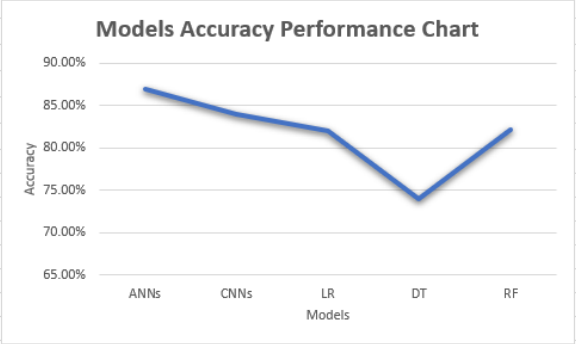

Fig. 2. Models Accuracy Metrics

2) Precision: Precision measures the proportion of true positive predictions among all the positive predictions made by the model. It is critical in medical diagnosis, especially for heart disease prediction, where false positives can lead to unnecessary treatments. Decision Trees and Random Forests exhibited perfect precision (1.00), meaning that these models made no false positive predictions. However, ANNs also performed well, with a precision of 0.91, showing its robustness in minimizing false positives. CNNs and LR had lower precision at 0.85 and 0.89, respectively, which might indicate a higher tendency for these models to produce false positives compared to tree-based models.

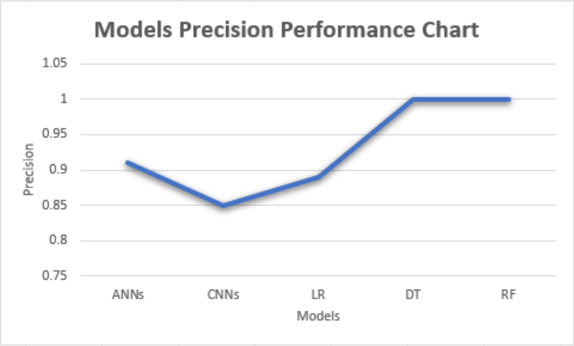

Fig. 3. Models Precision Metrics

3) *Recall:* Recall measures the model’s ability to correctly identify true positive cases, making it particularly important in medical diagnostics where it’s crucial to catch as many actual cases of heart disease as possible. Both Decision Trees and Random Forests achieved perfect recall (1.00), suggesting these models captured all true positive cases of heart disease. However, the ANN model achieved a slightly lower recall of 0.84, and Logistic Regression showed the lowest recall (0.79), meaning that these models might be missing a higher number of actual heart disease cases.

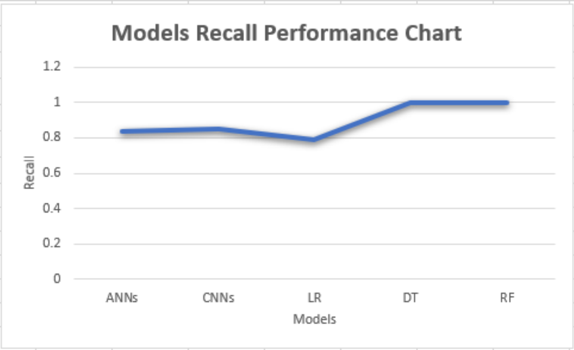

Fig. 4. Models Recalls Metrics

4) *F1-score:* The F1-score is the harmonic mean of precision and recall, offering a balanced assessment of a model's overall performance. Decision Trees and Random Forests achieved an F1-score of 1.00, indicating that these models excelled in both precision and recall. ANNs (0.87) and CNNs (0.85) demonstrated strong F1-scores as well, suggesting they offered a balanced performance. Logistic Regression, with an F1-score of 0.84, also performed reasonably well but lagged behind the other models in capturing an overall balance between precision and recall.

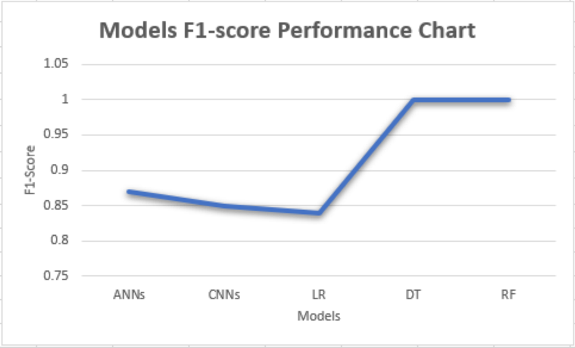

Fig. 5. Models FI-Score Metrics

B. *Discussion*

This study assessed the performance of several machine learning (ML) and deep learning (DL) models in predicting heart disease, emphasizing key performance metrics such as accuracy, precision, recall, and the F1-score. The models evaluated include artificial neural networks (ANNs), convolutional neural networks (CNNs), logistic regression (LR), decision trees (DT), and random forests (We thoroughly analysed each model's performance to identify its strengths and limitations in the context of heart disease prediction.

1) *Performance Based on Accuracy:* Accuracy measures the proportion of correctly classified instances among the total number of cases. The results indicate that ANNs achieved the highest accuracy at 87%, followed by CNNs with 84%, and Logistic Regression (LR) with 82%. Decision Trees (DT) had the lowest accuracy at 74%, while Random Forests (RF) performed similarly to Logistic Regression, with an accuracy of 82.1%. ANNs and CNNs, as deep learning models, demonstrated higher accuracy due to their ability to capture complex patterns in the data. However, the slight performance difference between ANNs and CNNs suggests that while CNNs excel at image data tasks, ANNs might be more suited for structured tabular data like medical records. Logistic Regression and Random Forests, despite their simplicity compared to deep learning models, still performed well with accuracy levels above 82%, indicating their utility in heart disease prediction. Decision Trees, while interpretable, exhibited lower accuracy, which could be due to overfitting on the training data, as they tend to create overly complex decision boundaries for imbalanced datasets.

2) *Precision, Recall, and F1-Score Analysis:*

a) *Precision:* Precision measures the proportion of true positive predictions out of all instances predicted as positive. ANNs had the highest precision at 0.91, indicating its effectiveness at reducing false positives. Logistic Regression followed with a precision of 0.89, while CNNs had a precision of 0.85. Decision Trees and Random Forests had perfect precision scores of 1.00, but this high score must be interpreted cautiously, as it may indicate that the models were not challenged with many false positives, possibly due to overfitting on certain classes. The high precision of ANNs underscores their ability to reduce false positives, which is critical in clinical settings to prevent healthy patients from being misdiagnosed with heart disease. While, decision Trees and Random Forests achieving perfect precision might suggest that they are overly biased toward the positive class, possibly sacrificing generalization for high precision.

b) *Recall***:** Recall, which measures the proportion of correctly predicted positive cases out of all actual positives, was similar across ANNs, CNNs, and Logistic Regression, with ANNs having the highest recall at 0.84, CNNs at 0.85, and Logistic Regression slightly lower at 0.79. Decision Trees and Random Forests both had recall values of 1.00, indicating perfect recall performance in this dataset. The high recall of ANNs and CNNs suggests that they are effective at identifying most patients with heart disease, minimizing the risk of false negatives (i.e., missed diagnoses). In addition, Decision Trees and Random Forests' perfect recall likely results from their tendency to classify most instances as positive, which can lead to overfitting. In clinical practice, models with high recall are preferred for disease detection to ensure that fewer cases are missed, but it is essential to balance this with precision to avoid excessive false positives.

c) *F1-Score:* The F1-score offers a balanced measure of precision and recall by calculating their harmonic mean. ANNs achieved the highest F1-score of 0.87, closely followed by CNNs at 0.85, and Logistic Regression at 0.84. Both Decision Trees and Random Forests reached perfect F1-scores of 1.00, likely due to overfitting.The high F1-score of ANNs reflects its ability to maintain an effective balance between precision and recall, making it a robust model for heart disease prediction, where minimizing both false positives and false negatives is essential. CNNs and Logistic Regression also achieved competitive F1-scores, further validating their suitability for predicting heart disease in scenarios where balanced performance is essential. The perfect F1-scores for Decision Trees and Random Forests, while impressive, likely indicate overfitting to the data, particularly in handling imbalanced datasets. This overfitting compromises the generalizability of the models to new data, which is a significant concern in clinical applications.

3) Model Strengths and Weaknesses: Each model exhibits unique strengths and weaknesses that influence its applicability in heart disease prediction.ANNs provided the best overall balance between accuracy, precision, recall, and F1-score, making it a robust choice for clinical decision support systems. However, its complexity may limit interpretability, which is crucial for clinicians. CNNs, while commonly used for image-based tasks, performed similarly to ANNs in this study, showing that they can also handle structured data effectively. The slightly lower performance compared to ANNs may reflect the need for further optimization. Logistic Regression demonstrated good accuracy and precision, making it a reliable choice for interpretable and transparent models. However, its slightly lower recall suggests it might miss some positive cases of heart disease, which is a drawback in a high-stakes environment like healthcare. Decision Trees and Random Forests, despite achieving perfect precision, recall, and F1-scores, are likely overfitting the training data. While their interpretability is an advantage, their lower generalizability poses a risk when applied to new patient data. Further optimization and regularization would be necessary to prevent overfitting.

4) Challenges and Opportunities: The challenges observed in this research stem from model complexity, interpretability, and the imbalanced nature of the dataset:

a) *Imbalanced Datasets:* Models like Decision Trees and Random Forests performed extremely well in metrics such as precision and recall, but this performance was likely due to overfitting. Addressing this imbalance using techniques such as oversampling, under sampling, or using class weights during model training could mitigate these issues.

b) *Interpretability vs. Performance*: While deep learning models (ANNs, CNNs) achieved high accuracy and F1-scores, their lack of interpretability can be a limitation in clinical settings where decision transparency is critical. Techniques like SHAP or LIME can be employed to make these models more interpretable.

c) *Generalization:* The need to ensure that models generalize well to new data is critical. Cross-validation techniques and regularization methods can help improve the robustness of the models, especially for Decision Trees and Random Forests.

V. CONCLUSION

Given the urgent need for precise, dependable, and understandable clinical decision support systems, this study emphasizes the significance of contrasting machine learning (ML) and deep learning (DL) models for heart disease prediction. The results demonstrate that computational intelligence models—in particular, Convolutional Neural Networks (CNNs) and Artificial Neural Networks (ANNs)—are useful tools for predicting cardiac disease because they can achieve high accuracy, precision, recall, and F1-scores. The study also highlights the trade-offs between interpretability and model complexity, especially for deep learning models, which are less transparent than more straightforward models like logistic regression.

The performance of Decision Trees and Random Forests, while perfect on some metrics, suggests overfitting, especially in the context of imbalanced datasets. This result disconfirms the assumption that high precision and recall necessarily indicate generalizability and trustworthiness. The findings support the need for further investigation into methods that address overfitting and ensure that models are robust when applied to new, unseen data.

At a broader level, these results have significant implications for clinical practice. The balance between accuracy and interpretability is crucial, as clinicians require models that are not only accurate but also explainable. The proposition that more complex models like ANNs and CNNs outperform traditional models such as Logistic Regression and Decision Trees is confirmed, though with the caveat that interpretability remains a challenge. This research suggests that future work should focus on improving model transparency, addressing dataset imbalances, and ensuring that these models can be trusted in real-world clinical settings where patient outcomes depend on accurate predictions.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my research supervisor for their invaluable guidance and support throughout this study. Special thanks to the whole staff of the Department of Computer Science, Faculty of Physical Sciences, University of Maiduguri for their support. Although no grants or financial support were received for this research, I deeply appreciate the encouragement and assistance from all those involved.

VI. REFERENCES

1] M. M. Ahsan and Z. Siddique, "Machine Learning-Based Heart Disease Diagnosis: A Systematic Literature Review," Artif Intell Med, vol. 128, Dec. 2021, doi: 10.1016/j.artmed.2022.102289.

[2] "Cardiovascular diseases." Accessed: Oct. 06, 2024. [Online]. Available: https://www.who.int/health-topics/cardiovascular-diseases#tab=tab_1

[3] C. M. Bhatt, P. Patel, T. Ghetia, and P. L. Mazzeo, "Effective Heart Disease Prediction Using Machine Learning Techniques," Algorithms 2023, Vol. 16, Page 88, vol. 16, no. 2, p. 88, Feb. 2023, doi: 10.3390/A16020088.

[4] R. Indrakumari, T. Poongodi, and S. R. Jena, "Heart Disease Prediction using Exploratory Data Analysis," Procedia Comput Sci, vol. 173, pp. 130–139, 2020, doi: 10.1016/J.PROCS.2020.06.017.

[5] S. Mohapatra et al., "A stacking classifiers model for detecting heart irregularities and predicting Cardiovascular Disease," Healthcare Analytics, vol. 3, Nov. 2023, doi: 10.1016/J.HEALTH.2022.100133.

[6] D. K. Plati et al., "A Machine Learning Approach for Chronic Heart Failure Diagnosis," Diagnostics 2021, Vol. 11, Page 1863, vol. 11, no. 10, p. 1863, Oct. 2021, doi: 10.3390/DIAGNOSTICS11101863.

[7] G. Tartarisco et al., "An intelligent Medical Cyber–Physical System to support heart valve disease screening and diagnosis," Expert Syst Appl, vol. 238, p. 121772, Mar. 2024, doi: 10.1016/J.ESWA.2023.121772.

[8] I. El Naqa and M. J. Murphy, "What Is Machine Learning?," Machine Learning in Radiation Oncology, pp. 3–11, 2015, doi: 10.1007/978-3-319-18305-3_1.

[9] R. Bhardwaj, … A. N.-2017 I. 41st annual, and undefined 2017, "A study of machine learning in healthcare," ieeexplore.ieee.orgR Bhardwaj, AR Nambiar, D Dutta2017 IEEE 41st annual computer software and applications, 2017•ieeexplore.ieee.org, Accessed: Oct. 06, 2024. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8029924/

[10] A. Ogunpola, F. Saeed, S. Basurra, A. M. Albarrak, and S. N. Qasem, “Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases,” Diagnostics 2024, Vol. 14, Page 144, vol. 14, no. 2, p. 144, Jan. 2024, doi: 10.3390/DIAGNOSTICS14020144.

Change these to IEEE format

[11] P. Rani et al., “An Extensive Review of Machine Learning and Deep Learning Techniques on Heart Disease Classification and Prediction,” Archives of Computational Methods in Engineering, vol. 31, no. 6, pp. 3331–3349, Aug. 2024, doi: 10.1007/S11831-024-10075-W/METRICS.

[12] Y. Rimal and N. Sharma, “Hyperparameter optimization: a comparative machine learning model analysis for enhanced heart disease prediction accuracy,” Multimedia Tools and Applications, vol. 83, no. 18, pp. 55091–55107, May 2024, doi: 10.1007/S11042-023-17273-X/METRICS.

[13] V. K. Sudha and D. Kumar, “Hybrid CNN and LSTM Network For Heart Disease Prediction,” SN Computer Science, vol. 4, no. 2, pp. 1–10, Mar. 2023, doi: 10.1007/S42979-022-01598-9/METRICS.

[14] M. Anand Kumar, K. Chandra Purohit, A. Singh, and S. Bhatt, “Heart Disease Prediction Model using Deep Learning Algorithms,” Webology, vol. 18, no. 4, pp. 1735–188, 2021, doi: 10.29121/WEB/V18I4/106.

[15] T. M. A. Monisha Sharean and G. Johncy, “Deep learning models on Heart Disease Estimation - A review,” Journal of Artificial Intelligence and Capsule Networks, vol. 4, no. 2, pp. 122–130, Jul. 2022, doi: 10.36548/JAICN.2022.2.004.

[16] C. Zhou et al., “A comprehensive review of deep learning-based models for heart disease prediction,” Artificial Intelligence Review, vol. 57, no. 9, pp. 1–50, Sep. 2024, doi: 10.1007/S10462-024-10899-9/TABLES/10.

[17] M. Mandava and S. Reddy Vinta, “MDensNet201-IDRSRNet: Efficient cardiovascular disease prediction system using hybrid deep learning,” Biomedical Signal Processing and Control, vol. 93, p. 106147, Jul. 2024, doi: 10.1016/J.BSPC.2024.106147.

[18] G. A. Ansari, S. S. Bhat, M. D. Ansari, S. Ahmad, J. Nazeer, and A. E. M. Eljialy, “Performance Evaluation of Machine Learning Techniques (MLT) for Heart Disease Prediction,” Computational and Mathematical Methods in Medicine, vol. 2023, 2023, doi: 10.1155/2023/8191261.

[19] K. Drożdż et al., “Risk factors for cardiovascular disease in patients with metabolic-associated fatty liver disease: a machine learning approach,” Cardiovascular Diabetology, vol. 21, no. 1, Dec. 2022, doi: 10.1186/S12933-022-01672-9.

[20] F. S. Alotaibi, “Implementation of machine learning model to predict heart failure disease,” International Journal of Advanced Computer Science and Applications, vol. 10, no. 6, pp. 261–268, 2019, doi: 10.14569/IJACSA.2019.0100637.

[21] V. Sharma, S. Yadav, and M. Gupta, “Heart Disease Prediction using Machine Learning Techniques,” in Proceedings of the IEEE 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Dec. 2020, pp. 177–181, doi: 10.1109/ICACCCN51052.2020.9362842.

[22] Q. Bi, K. E. Goodman, J. Kaminsky, and J. Lessler, “What is Machine Learning? A Primer for the Epidemiologist,” American Journal of Epidemiology, vol. 188, no. 12, pp. 2222–2239, Dec. 2019, doi: 10.1093/AJE/KWZ189.