International Journal of AI, BigData, Computational and Management Studies
Noble Scholar Research Group | Volume 6, Issue 4, PP 217-224, 2025
ISSN: 3050-9416 | https://doi.org/10.63282/3050-9416.IJAIBDCMS-V6I4P125

Original Article

Comparative Study of Three Artificial Intelligence Algorithms for the Prediction and Classification of Cardiovascular Diseases from ECG Signals: Application to Chadian Patients

ALI Ouchar Cherif[1], RONGAR Octave[2], Conrad Onesime Oboulhas Tsahat[3], Khadidja Ousman Kossi[4], Aristide Mankiti Fati[5]
[1]*INSTA.*
[2]*University of Moundou*
[3]*ENASTIC*
[4,5]*Marien Ngouabi University*

Received On: 16/10/2025 ***Revised On: 17/11/2025*** ***Accepted On: 30/11/2025*** ***Published On: 10/12/2025***

Abstract: *In this paper, a comparative study of three artificial intelligence algorithms: Artificial Neural Networks (ANN), k-Nearest Neighbors (KNN), and Logistic Regression (LR) is conducted for the prediction and classification of cardiovascular diseases (CVD) based on electrocardiogram (ECG) data from Chadian patients.*
The experimental analysis relies on a dataset of 1,025 patients, comprising 14 clinical and electrocardiographic attributes, collected from local hospital databases.
Model performance was evaluated using standard metrics such as accuracy, precision, recall (sensitivity), and F1-score. The results demonstrate that the ANN model outperforms the other approaches, achieving an accuracy of 99.6%, a precision of 100%, a recall of 99.4%, and an F1-score of 99.6%, compared to slightly lower performance obtained with the KNN and LR models.
These findings confirm the strong potential of artificial neural networks for ECG signal analysis and decision support in the diagnosis of cardiovascular diseases. In the Chadian context, characterized by limited medical and technological resources, this approach represents a promising solution for early detection of CVD and improved patient management. However, further studies are required to validate the robustness of the proposed model on larger datasets and to integrate additional clinical risk factors.

Keywords: *Electrocardiogram (ECG), Artificial Neural Networks (ANN), k-Nearest Neighbors (KNN), Logistic Regression (LR), Artificial Intelligence, Cardiovascular Diseases, Prediction.*

1. Introduction

Cardiovascular diseases (CVDs) represent the leading cause of mortality worldwide, accounting for approximately 32% of all annual deaths and resulting in nearly 19.8 million deaths each year. This burden is particularly severe in low- and middle-income countries, where access to healthcare services and preventive programs remains limited [1].

The electrocardiogram (ECG) is a widely used non-invasive technique for recording the electrical activity of the heart and identifying various cardiac abnormalities, such as arrhythmias or signs of myocardial infarction [2].

With the growing volume of available data and the increasing complexity of ECG signals, machine learning algorithms have demonstrated strong potential for enhancing signal analysis and improving the prediction of cardiovascular diseases. Among these approaches, Artificial Neural Networks (ANN), Logistic Regression (LR), and the k-Nearest Neighbors (KNN) algorithm are frequently employed for medical data classification and have shown relevant performance in this domain [3].

The objective of this study is to integrate these technologies in order to predict and classify cardiovascular diseases based on collected ECG signals, to evaluate the performance of the selected models, and to discuss their practical implications for medical diagnosis, particularly within the Chadian context.

1.1. Related Works and State of the Art

The state-of-the-art highlights recent advances in artificial intelligence (AI) for the diagnosis of cardiovascular diseases using electrocardiogram (ECG) signals. Numerous studies have demonstrated that exploiting ECG data through machine learning algorithms significantly improves early detection, classification, and prediction of cardiac pathologies.

Several AI algorithms, including Artificial Neural Networks (ANN), Logistic Regression (LR), and the k-

Nearest Neighbors (KNN) algorithm, have been widely applied to the classification of cardiovascular abnormalities. Existing studies indicate that supervised learning and deep learning (DL)-based approaches achieve high performance levels, with accuracy rates generally ranging from 92.5% [7] to 98% [8], depending on the methods and datasets used. For instance, some works report accuracies as high as 99.6% for ANN-based models, outperforming KNN-based approaches (98.7%) [10] and logistic regression models [3]. These findings confirm the strong potential of AI combined with ECG signal analysis as a promising solution for the rapid detection and management of cardiovascular diseases.

However, undertaking in-depth research on the prediction of cardiovascular diseases using artificial intelligence would be inappropriate without considering prior studies that have contributed to the development of this field. Although the present work focuses on the implementation of AI- and ECG-based prediction and classification techniques in the specific context of Chadian patients, it is essential to situate this approach within the continuity of existing scientific contributions.

Cardiovascular diseases remain one of the leading causes of mortality worldwide, and Chad is no exception to this reality. In this context, early detection and reliable classification of cardiac pathologies emerge as critical levers for improving medical care. The combination of electrocardiographic (ECG) signals with artificial intelligence algorithms therefore constitutes an innovative and promising approach for predicting and classifying cardiovascular abnormalities.

In this regard, Asmaa Ameen et al. (2024) conducted a comprehensive study reviewing 82 research works focused on the use of ECG data for cardiovascular disease classification. Their findings indicate that deep learning methods, particularly convolutional neural networks (CNNs), significantly outperform traditional approaches such as support vector machines (SVMs) and random forests. This methodological superiority suggests strong potential for improving diagnostic accuracy, provided that datasets adapted to the local context particularly that of Chad are available.

Similarly, Zhao (2023) highlighted the effectiveness of transformer-based models, which are capable of capturing complex temporal relationships within ECG signals. These models demonstrate robustness when dealing with long or noisy signals, representing a notable advantage in resource-constrained clinical environments often characterized by outdated equipment or electromagnetic interference [2].

Furthermore, Maleki and Haeri (2024) proposed a wavelet transform-based approach to extract discriminative features from ECG signals across different frequency bands.

This method showed significant improvements in the automatic detection of cardiovascular anomalies, offering a potentially suitable solution for Chadian hospital structures, where the availability of specialized cardiology personnel remains limited [4].

In addition, Joloudari et al. (2021) confirmed the superiority of deep convolutional neural networks (DCNNs) over traditional methods in the detection of myocardial infarction. The adoption of such approaches could contribute to faster identification of critical cases and to a reduction in mortality associated with diagnostic delays [3,4].

Finally, Le et al. (2022) developed a lightweight and explainable deep learning system, named *LightX3ECG*, compatible with wearable devices. This solution, both efficient and resource-efficient, opens promising perspectives for continuous patient monitoring, particularly in rural or underserved areas, while reducing constraints related to hospital infrastructure [4].

Despite these significant advances, several cross-cutting limitations persist. First, the majority of existing studies rely on datasets originating from developed countries, limiting their representativeness for the Chadian context. Second, advanced models such as CNNs or transformers require substantial computational resources, which are often difficult to access in local healthcare facilities. Finally, the lack of explainability of certain AI models frequently perceived as "black boxes" remains a major barrier to their adoption by healthcare practitioners.

Within this context, the present article aims to conduct a comparative study of prediction and classification methods for cardiovascular diseases using artificial intelligence algorithms and ECG data, while taking into account the constraints and specific characteristics of Chadian patients.

2. Methodology

The objective of this study is to exploit electrocardiogram (ECG) data and to compare three machine learning algorithms Artificial Neural Networks (ANN), Logistic Regression (LR), and k-Nearest Neighbors (KNN) in order to identify the most effective model for predicting cardiovascular diseases (CVD). To achieve this objective, several methodological steps were followed.

2.1. Data Collection

Data collection is based on ECG recordings from patients with and without cardiovascular diseases, obtained from the databases of Chadian hospitals. This database, initially established in 1988 and updated on May 28, 2025, contains 76 attributes, including a predictive attribute indicating the presence or absence of cardiac pathology.

Table 1: Patient Cardiovascular Data and Heart Disease Indicators

Index	Age	Sex	TAR	BPS	Chol	Gly	ECG	FCMA	Exang	Dep	Slope	NNP	Thal	Cardio
0	52	1	0	125	212	0	1	168	0	1.0	2	2	3	0

1	53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
2	70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
3	61	1	0	148	203	0	1	161	0	0.0	2	1	3	0
4	62	0	0	138	294	1	1	106	0	1.9	1	3	2	0
…	…	…	…	…	…	…	…	…	…	…	…	…	…	…
1024	54	1	0	120	188	0	1	113	0	1.4	1	1	3	0

However, in accordance with common practices adopted in the scientific literature, a subset of 14 clinical and electrocardiographic attributes was selected. This subset includes the target variable *Cardio*, encoded as follows:

Legend :

- Sex: 1 = Male, 0 = Female
- Heart Disease (Cardio): 1 = Presence, 0 = Absence
- Chest Pain Type (CP):
  - 0: Typical angina
  - 1: Atypical angina
  - 2: Non-anginal pain
  - 3: Asymptomatic

In this study, the dataset used consists of 1,025 patients and 14 attributes, derived exclusively from this database.

2.2. Data and Attribute Description

Table 2 presents the clinical and electrocardiographic characteristics used in this study, along with their descriptions and data types.

Table 2: Clinical Characteristics of Patients (Ministry of Public Health of Chad)

No.	Attribute	Description	Type
1	Age	Patient age (29–77 years)	Numerical
2	Sex	Patient sex (0: female, 1: male)	Categorical
3	CP	Chest pain type (0: typical angina, 1: atypical angina, 2: non-anginal pain, 3: asymptomatic)	Categorical
4	Trestbps	Resting blood pressure (mm Hg)	Numerical
5	Chol	Serum cholesterol (mg/dL, 126–564)	Numerical
6	FBS	Fasting blood sugar > 120 mg/dL (1: true, 0: false)	Categorical
7	Restecg	Resting ECG results (0–2)	Categorical
8	MaxHR	Maximum heart rate achieved (71–202)	Numerical
9	Exang	Exercise-induced angina (1: yes, 0: no)	Categorical
10	ST Depression	ST segment depression induced by exercise relative to rest	Numerical
11	Slope	Slope of the ST segment at peak exercise	Categorical
12	CA	Number of major vessels (0–3)	Numerical
13	Thal	Blood disorder (thalassemia)	Categorical
14	Cardio	Presence or absence of heart disease	Categorical

2.3. Data Preprocessing

Before model training, the ECG data underwent an essential preprocessing phase to enhance signal quality and ensure the reliability of the results. This phase included:

- removal of artifacts and noise present in ECG signals,
- extraction of relevant features such as heart rate, PR and QT intervals, and the amplitude of specific ECG waves,
- normalization of data to ensure a homogeneous scale across different variables.

Data cleaning and preparation were performed using dedicated Python libraries, including NumPy, Pandas, Matplotlib, Seaborn, Graphviz, and TensorFlow. The final dataset was then structured in the form of a DataFrame to facilitate analysis and machine learning processes.

2.4. Model Training

ANN, KNN, and LR algorithms were employed to build classification models capable of predicting the presence or absence of cardiovascular diseases based on the selected ECG features.

The dataset was divided into two subsets:

- a **training set**, used to train the models,
- a **test set**, used to evaluate model performance on unseen data.

Each model was trained using the ECG features and their corresponding labels, enabling an objective comparison of their generalization ability, accuracy, and robustness in predicting cardiovascular diseases.

2.5. Machine Learning Models

2.5.1. Artificial Neural Networks (ANN)

An Artificial Neural Network (ANN) is a machine learning model inspired by the structure of biological neural networks. It consists of artificial neurons organized into layers (input, hidden, and output) that enable the processing of complex and non-linear data. The connections between neurons are weighted and iteratively adjusted through a

learning process aimed at minimizing a cost function. In deep learning architectures, these networks include multiple hidden layers, allowing the extraction of hierarchical data representations, which is a major advantage for the analysis of raw or transformed ECG signals [3].

ANNs are widely used for the automatic classification of ECG signals, as they can learn directly from raw or preprocessed inputs without requiring manual feature extraction. Recent studies in the literature demonstrate that neural network-based architectures, including deep variants, significantly improve the accuracy of ECG-based diagnostic systems [9].

2.5.2. k-Nearest Neighbors (KNN) Algorithm

The k-Nearest Neighbors (KNN) algorithm is a non-parametric supervised learning method primarily used for classification tasks. The fundamental principle of KNN is based on the assumption that data points that are close to each other in the feature space are likely to belong to the same class. For a new sample, the algorithm identifies the *k* nearest observations in the training dataset and assigns the class that is most frequent among these neighbors [3].

KNN does not require an explicit training phase, as most computations are performed during the classification stage, which makes it simple to understand and implement. However, its performance can be negatively affected by high-dimensional data or class imbalance. Moreover, the optimal choice of *k* (number of neighbors) directly influences the generalization capability of the model in complex contexts such as ECG signal analysis [10].

2.5.3. Logistic Regression (LR)

Logistic Regression (LR) is a statistical technique used to model the probability that an observation belongs to a binary class (e.g., presence or absence of cardiovascular disease). Unlike linear regression, LR employs the logistic (sigmoid) function to constrain predictions within the range [0, 1], making it well suited for binary classification problems [3].

The model parameters are estimated using maximum likelihood estimation, which provides a probabilistic interpretation of the results. In medical applications, logistic regression is valued for its simplicity, computational efficiency, and interpretability, as it enables the estimation of the impact of each explanatory variable (e.g., age, heart rate) on the probability of disease occurrence [10].

3. Tools and Experimental Environment

This study was conducted in a computational environment combining open-source software tools and cloud-based computing platforms, adapted to biomedical data analysis and the implementation of artificial intelligence algorithms. All processing steps—from ECG signal preprocessing to model training and evaluation—were carried out within a reproducible framework consistent with scientific publication standards.

3.1. Hardware Environment

The experiments were performed using the Google Colaboratory (Google Colab) platform, which allows the execution of Python notebooks without prior local configuration and provides access to CPU resources, as well as graphical processing units (GPUs) when required [6]. This solution is particularly suitable for environments with limited hardware resources, such as those commonly encountered in Chadian hospital infrastructures.

3.2. Software Environment

The software environment was primarily based on the Python programming language, which is widely used in artificial intelligence and medical data analysis. The experiments were developed and executed using Jupyter Notebook, an interactive environment that integrates code, visualizations, and explanatory text, thereby ensuring experiment reproducibility [6].

3.3. Data Processing and Analysis Libraries

ECG data preprocessing, manipulation, and analysis were carried out using standard scientific libraries. NumPy was used for numerical computations, while Pandas facilitated data structuring, cleaning, and management in the form of DataFrames [12].

Statistical operations and certain signal transformations were supported by SciPy, whereas Matplotlib was employed for generating figures, plots, and visualizations of model performance.

3.4. Machine Learning and Deep Learning Libraries

Machine learning models were implemented using specialized libraries. Scikit-learn was used for implementing Logistic Regression (LR) and k-Nearest Neighbors (KNN) algorithms, as well as for splitting the dataset into training and testing sets and computing evaluation metrics.

The Artificial Neural Network (ANN) model was developed using TensorFlow, an open-source deep learning framework [16], with Keras as a high-level API facilitating the design, training, and optimization of neural networks [12].

3.5. Performance Evaluation and Visualization

Model evaluation was conducted using confusion matrices and standard classification metrics, including accuracy, precision, recall (sensitivity), and F1-score.

The learning dynamics of the ANN model were analyzed through accuracy and loss curves, enabling the identification of convergence behavior and potential overfitting.

4. Performance Evaluation Criteria for Classification Models

Performance evaluation aims to measure a model's ability to generalize its predictions to data that were not used during the training phase. This step is essential for estimating model behavior under real-world conditions. In this study,

the evaluation of classification model performance primarily relies on the analysis of the confusion matrix, from which several standard metrics are derived.

4.1. Confusion Matrix

The confusion matrix is a table that summarizes the prediction results of a classification model by comparing the true values of the target variable with the values predicted by the model [11]. It highlights correct and incorrect predictions, distributed across four categories:

- True Positive (TP): the prediction is positive and the actual value is also positive (e.g., a diseased patient correctly identified as diseased);
- True Negative (TN): the prediction is negative and the actual value is negative (e.g., a healthy patient correctly identified as healthy);
- False Positive (FP): the prediction is positive while the actual value is negative (e.g., a healthy patient incorrectly classified as diseased);
- False Negative (FN): the prediction is negative while the actual value is positive (e.g., a diseased patient incorrectly classified as healthy).

The confusion matrix forms the basis for computing the main evaluation metrics, including accuracy, precision, recall (sensitivity), and the F1-score [11].

Table 3: *Confusion Matrix*

Actual Class \ Predicted Class	**Positive**	**Negative**
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

4.2. Accuracy

Accuracy represents the proportion of correct predictions relative to the total number of predictions made [11]. It is defined as follows:

$$\textbf{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \ (1) \qquad (1)$$

4.3. Recall (Sensitivity)

Recall, also referred to as sensitivity, measures the ability of the model to correctly identify positive instances. It corresponds to the ratio between the number of true positives and the total number of actual positive cases [11]:

$$\textbf{Recall} = \frac{TP}{TP+FN} \qquad (2)$$

4.4. Precision

Precision evaluates the proportion of observations correctly classified as positive among all observations predicted as positive [11]. It is expressed as:

$$\textbf{Precision} = \frac{TP}{TP+FP} \qquad (3)$$

4.5. F1-score

The F1-score is a composite metric that combines precision and recall. It is particularly suitable for classification problems involving imbalanced datasets [11]. The F1-score corresponds to the harmonic mean of precision and recall and is defined as:

$$\textbf{F1 - score} = \frac{2*Precision*Recall}{Precision+Recall} = \frac{2*TP}{2*TP+FP+FN} \qquad (4)$$

5. Evaluation of Algorithm Performance and Model Selection

This section presents the results obtained by applying different classification models to the dataset used for predicting cardiovascular diseases.

5.1. Accuracy

The histogram presented illustrates the accuracy values obtained for the different algorithms evaluated in this study.

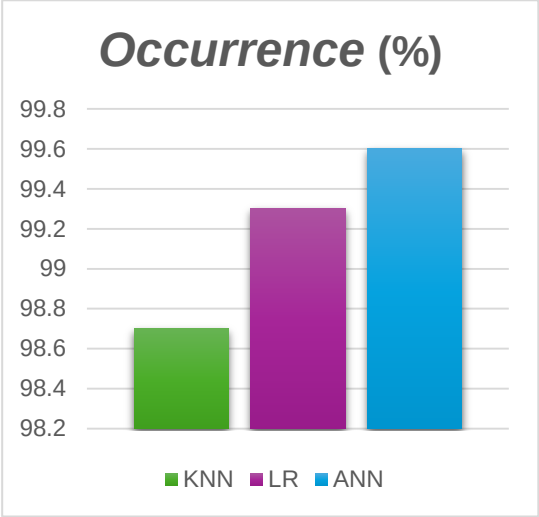

Fig 1: Accuracy Results of the Evaluated Models

Figure 1 shows that, although the overall performances of the models are close, the ANN algorithm achieves the highest accuracy on the dataset used.

5.2. Precision

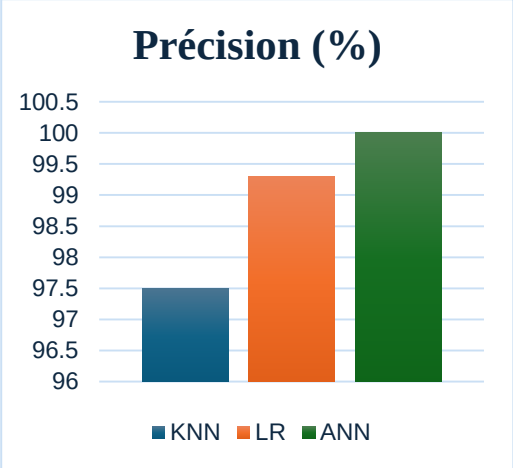

Fig 2: Results Obtained in Terms of Precision

Figure 3 compares the performance of the different models in terms of precision. The results indicate that the ANN model achieves a precision of 100%, demonstrating its ability to correctly identify patients with cardiovascular diseases. The KNN and LR models also exhibit good performance, but remain slightly inferior to that of the ANN.

5.3. Sensitivity

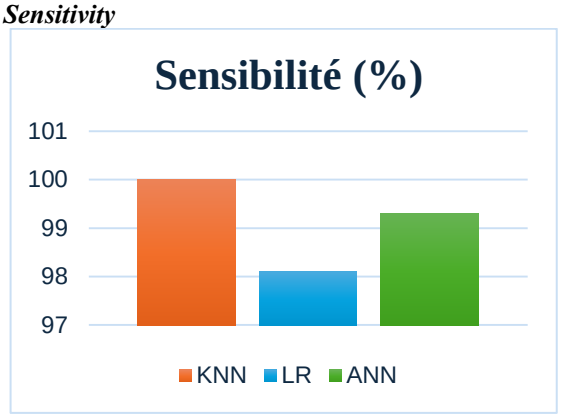

Fig 3: Results Obtained in Terms of Sensitivity

Figure 4 highlights the comparison of the algorithms in terms of sensitivity. The results show that the KNN algorithm achieves the highest sensitivity, indicating a strong capability to detect positive cases of cardiovascular diseases.

5.4. F1-score

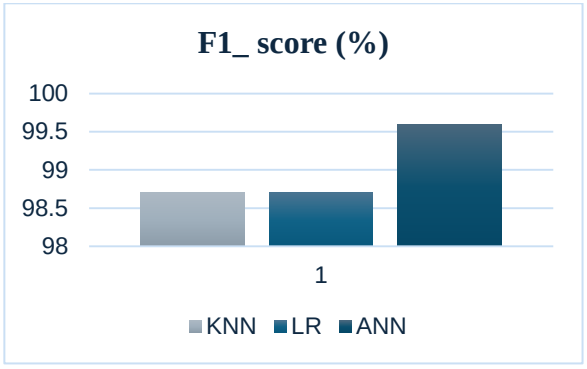

Fig 4: Results Obtained in Terms of F1-score

Figure 4 presents a comparison of the models based on the F1-score. It appears that the ANN algorithm provides the best trade-off between precision and recall, achieving the highest F1-score among the evaluated models.

6. Experimental Results and Comparative Analysis

6.1. Model Selection

Based on the experiments conducted using the Artificial Neural Network (ANN), k-Nearest Neighbors (KNN), and Logistic Regression (LR) models, it appears that all three approaches provide overall satisfactory performance for the classification of cardiovascular diseases. The selection of these models is motivated by their demonstrated effectiveness in the literature for similar biomedical data classification tasks. The obtained results show that each model is capable of generalizing adequately and distinguishing patients with cardiovascular diseases from healthy individuals.

However, among the three evaluated algorithms, the ANN proved to be the most performant. It achieved a precision of 100% and an accuracy of 99.6%, correctly classifying 303 out of 308 observations. The KNN model, although performant, exhibited slightly lower results, with a precision of 99.5% and an accuracy of 99.3%. This difference can be attributed to the non-parametric nature of KNN and its sensitivity to data distribution. Finally, the LR model achieved a precision of 97.5% and an accuracy of 98.7%, confirming its role as a robust baseline model, albeit less effective than non-linear approaches.

After training the KNN, LR, and ANN models, the test dataset was used to predict class labels and to select the optimal model. This selection was based on a joint analysis of several performance metrics, including accuracy, precision, sensitivity (recall), and F1-score. In this study, accuracy was prioritized as the primary criterion, given the balanced nature of the dataset used.

The results confirm that the ANN provides more reliable and stable predictions than the other models. **Table 3** presents a summary of the performance evaluation metrics derived from the confusion matrices for each algorithm.

Table 4: Performance Evaluation Results of the Models

Model	Accuracy (%)	Precision (%)	Sensitivity (%)	F1-score (%)
KNN	99.3	99.4	100.0	98.6
LR	97.5	97.5	98.2	98.6
ANN	99.6	100.0	99.4	99.6

6.2. In-depth Discussion of the Selected ANN Model Results

6.2.1. Results

6.2.1.1. Distribution of Cardiovascular Diseases by Sex

The analysis of the distribution of cardiovascular diseases by sex shows a slightly higher prevalence among women compared to men. However, the overall distribution remains relatively balanced between the two sexes, suggesting that sex, when considered in isolation, is not a major discriminative factor within the studied dataset.

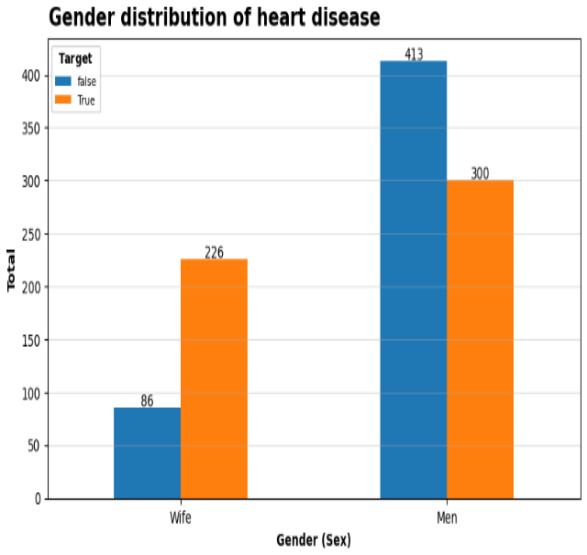

Fig 5: Distribution of Heart Diseases by Sex

6.2.1.2. Analysis of Cardiovascular Diseases as a Function of Age

The age-based scatter plot indicates that the majority of patients, whether affected or not by cardiovascular diseases, fall within the 50–70 age range. It also appears that patients with cardiovascular diseases generally exhibit a higher heart rate than those without cardiac pathology, highlighting the importance of age and certain physiological parameters in the prediction of heart diseases.

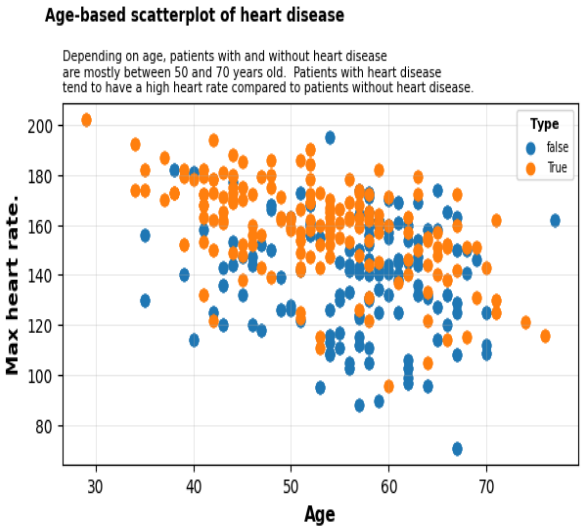

Fig 6: Scatter Plot of Cardiovascular Diseases as a Function of Age

6.2.2. Analysis of the Strengths and Limitations of the Proposed Models

The analysis of the learning curves of the ANN model highlights a rapid and parallel increase in both training and validation accuracy during the initial epochs, reflecting the model's strong ability to learn relevant features present in the data. Beyond a certain threshold, these accuracies stabilize, indicating that the model has reached an optimal learning level. The validation accuracy reaches a maximum of 99.6% at epoch 500.

In addition, the training and validation loss curves show a gradual and parallel decrease, demonstrating a continuous improvement in model performance. The stabilization of the loss values after a certain number of epochs suggests that the ANN has achieved a balance between learning and generalization, with no significant signs of overfitting.

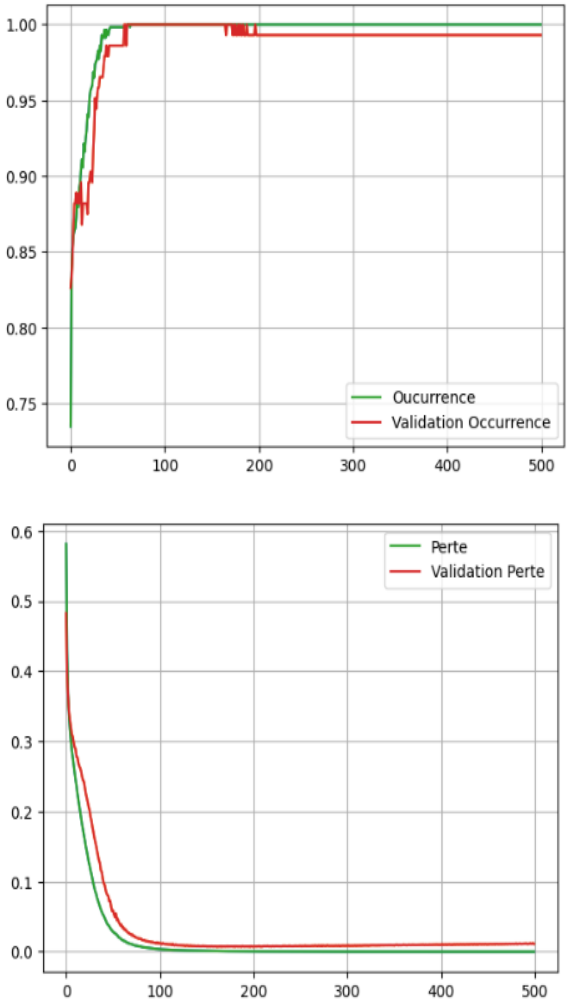

Fig 7: Accuracy and Loss Curves over Training Epochs

In this study, the dataset was split into 70% for training and 30% for testing. The ANN model was evaluated on 308 test samples, among which 303 were correctly classified, corresponding to an accuracy of 98.37%. The five observed misclassifications represent an error rate of 1.62%.

Overall, these results demonstrate that the use of ECG recordings combined with the ANN algorithm yields particularly promising predictive performance for cardiovascular diseases. The model is able to effectively identify specific patterns and features associated with different cardiac pathologies, thereby improving the reliability and accuracy of predictions.

7. Discussion

The results obtained in this study are generally consistent with international research on the application of artificial intelligence to the prediction of cardiovascular diseases using electrocardiographic signals. However, a critical analysis highlights both methodological convergences, performance differences, and contextual specificities that distinguish our approach.

Recent studies reported in the literature [7] and [8] confirm that machine learning algorithms, particularly neural networks, achieve high performance in the classification of cardiovascular diseases. These works emphasize the ability of non-linear models to extract complex relationships from ECG data. Nevertheless, most of these studies rely on datasets originating from developed countries, such as the UCI or Cleveland datasets, which limits their direct transferability to different healthcare contexts.

In this respect, the main contribution of our study lies in the exploitation of local Chadian data, enabling the evaluation of model robustness in an environment that is rarely explored in the international literature. While study [7] reports an accuracy of 95.71% for an ANN model applied to the UCI dataset, and other works [7], [8] report performance levels ranging from 75% to 92.86% depending on the algorithms used, our approach achieves a precision of 100% and an accuracy of 99.6%. This performance gap suggests that adapting models to local demographic and clinical specificities can play a decisive role in improving predictive outcomes.

Moreover, unlike some international studies that favor complex architectures such as deep convolutional neural networks (CNNs) or transformer-based models [2], our work demonstrates that a classical artificial neural network, when properly parameterized and trained on relevant data, can achieve comparable or even superior performance. This finding is particularly significant for resource-constrained environments, where the computational power required to train deep models is often not readily available.

In addition, several studies in the literature highlight the limitations of advanced models in terms of interpretability, which are often perceived as “black boxes” by healthcare practitioners. In contrast, the comparative analysis conducted in this study between ANN, KNN, and LR offers a balance between predictive performance and implementation simplicity, thereby promoting more realistic adoption within local healthcare structures.

It should nevertheless be emphasized that, as in many international studies, the high performance observed may be influenced by sample size, feature selection, and the balanced nature of the dataset used. In this regard, the integration of a larger volume of data, as well as the inclusion of additional clinical risk factors, could further strengthen the results and enhance the generalization capability of the model.

Overall, this study confirms the major conclusions of the international literature regarding the effectiveness of artificial intelligence algorithms for cardiovascular disease prediction, while highlighting the importance of contextual adaptation of models. The performance achieved by the ANN using Chadian data suggests that AI-based solutions, when properly calibrated, can constitute reliable and accessible decision-support tools, even in resource-limited environments.

8. Conclusion

This article proposed an effective approach for the prediction and classification of cardiovascular diseases, based on the exploitation of electrocardiogram (ECG) signals and the use of the Artificial Neural Network (ANN) algorithm. The study relied on a dataset of 1,025 clinical records, comprising 14 relevant attributes extracted from hospital databases. Model performance was evaluated using standard metrics, including accuracy, precision, sensitivity (recall), and F1-score.

The obtained results are particularly compelling, with an accuracy of 99.6%, a precision of 100%, an F1-score of 99.6%, and a sensitivity of 99.4%. These performance levels confirm the ANN model’s ability to generalize effectively and to reliably distinguish patients with cardiovascular diseases from healthy subjects. They also highlight the value of combining ECG data with machine learning techniques for medical decision support.

In a context such as Chad, characterized by limited medical and technological resources, this approach represents a promising solution for assisting healthcare professionals in the early detection of cardiovascular pathologies. In this regard, a relevant development perspective lies in the implementation of a web-based prediction platform incorporating an alert system intended for physicians and, ultimately, for patients.

The combination of ECG data and ANN algorithms thus opens encouraging prospects for earlier medical intervention, potentially improving clinical outcomes and reducing morbidity and mortality associated with cardiovascular diseases. Nevertheless, further work remains necessary to consolidate model robustness, integrate additional clinical risk factors, and validate performance on larger and more diverse datasets.

In conclusion, this study demonstrates that the integration of ECG sensors and artificial neural network algorithms represents a significant technological advancement, capable of contributing sustainably to the improvement of early cardiovascular disease detection and the optimization of clinical outcomes, particularly in resource-limited settings. Although the achieved performance is high, it should be interpreted in light of the sample size and the specific context of the data used.

References

1. Organisation mondiale de la santé (OMS), *Maladies cardiovasculaires (MCV)*, 2017. [En ligne]. Disponible : https://www.who.int/fr/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)
2. M. Maleki and F. Haeri, “Identification of cardiovascular diseases through ECG classification using wavelet transformation,” *arXiv preprint*, arXiv:2024.xxxxx, 2024.
3. J. H. Joloudari, S. Mojrian, I. Nodehi, A. Mashmool, R. Alizadehsani, A. Mosavi, and U. R. Acharya, “Application of artificial intelligence techniques for

automated detection of myocardial infarction: A review," *arXiv preprint*, arXiv:2101, 2021.

4. K. H. Le, H. H. Pham, T. B. T. Nguyen, T. A. Nguyen, T. N. Thanh, and C. D. Do, "LightX3ECG: A lightweight and explainable deep learning system for 3-lead electrocardiogram classification," *arXiv preprint*, arXiv:2203, 2022.
5. Z. Zhao, "Transformer-based deep learning models for ECG diagnosis: An in-depth review," *arXiv preprint*, arXiv:2305, 2023.
6. H. Terghini, *Conception et réalisation d'une approche de deep learning pour l'Internet des objets (IoT)*, Mémoire/Thèse, 2020.
7. F. Ghebouli and A. Lakroum, *Application pour la classification et la migration entre classes de la maladie cardiaque*, Thèse de doctorat, Université Mohamed El Bachir El Ibrahimi, Bordj Bou Arréridj, Algérie, 2021.
8. W. Caesarendra, T. A. Hishamuddin, D. T. C. Lai, A. Husaini, L. Nurhasanah, A. Glowacz, and G. A. F. Alfarisy, "An embedded system using convolutional neural network model for online and real-time ECG signal classification and prediction," *Diagnostics*, vol. 12, no. 4, p. 795, 2022.
9. M. Latreche, *Classification des signaux électrocardiographiques (ECG) par les réseaux de neurones artificiels*, Mémoire/Thèse, 2021.
10. L. A. Behlouli and H. Hamdi, *Une approche hybride AG–KNN pour la prédiction des maladies cardiaques*, Thèse de doctorat, Université Akli Mohand Oulhadj–Bouira, Algérie, 2022.
11. Mhammedi, I. Yakoub, A. Ouahab, *et al.*, *La détection par apprentissage profond (Deep Learning) et métriques d'évaluation*, Thèse de doctorat, Université Ahmed Draïa – Adrar, Algérie, 2021.
12. P. Virtanen *et al.*, "SciPy 1.0: Fundamental algorithms for scientific computing in Python," *Nature Methods*, vol. 17, no. 3, pp. 261–272, 2020.
13. Ameen, I. Eldesouky Fattoh, T. Abd El-Hafeez and K. Ahmed, "*Advances in ECG and PCG-based cardiovascular disease classification : A review of deep learning and machine learning methods,"* Journal of Big Data, vol. 11, no. 1, p. 159, 2024, doi:10.1186/s40537-024-01011-7.