International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 5, pp 65-70, May-2025

https://doi.org/10.47001/IRJIET/2025.905008

Normalized Clinical Feature Neural Net (NCF-NN) for Cardiovascular Prediction

[1]Ruaa H. Ali Al-Mallah, [2]Marwa Mawfaq Mohamedsheet Al-Hatab

[1]Technical Engineering College for Computer and Artificial Intelligence, Northern Technical University, Mosul, Iraq

[2]Technical Engineering College, Northern Technical University, Mosul, Iraq

Authors E-mail: [1]ruaa_almallah@ntu.edu.iq, [2]marwa.alhatab@ntu.edu.iq

***Abstract* - Heart disease continues to be one of the leading causes of demise around the world, emphasizing the urgent need for effective early detection mechanisms. Normalized Clinical Feature Neural Net (NCF-NN), a neural network-based technique designed to categorize patients based on the likelihood of cardiovascular issues utilizing 13 clinical characteristics, is proposed in this research. The architecture involves two concealed layers with ReLU activation and L2 regularization, optimized using stochastic gradient descent and binary cross-entropy loss. Leveraging a dataset of 13 standardized clinical attributes extraction, the model attained a prediction accuracy of 98%, an AUC of 0.99, and consistently robust outcomes across other evaluation metrics. These conclusions underscore the model's potential as a practical diagnostic support tool in clinical environments, offering dependable risk prediction and contributing to more informed and proactive cardiovascular care. Separately, some patients exhibited multiple risk factors increasing the complexity of analysis, while others presented with only one or two characteristics highlighting the variance in presentations. The model successfully classified cases along this spectrum demonstrating its ability to evaluate diverse patient profiles.**

Keywords: Normalized Clinical Feature Neural Net (NCF-NN), cardiovascular disease, Machine learning, classification.

I. Introduction

Cardiovascular disease (CVD), more commonly referred to as heart disease, remains a top cause of morbidity and mortality globally, with millions of deaths attributed to it annually. Identifying Heart-Related Conditions In on-time manner helps reduce long-term complications and enhances the chances of survival in patients. Conventional diagnostics usually depend on fault serious interpretation of symptoms, clinical expertise, and invasive checks, which can be labor intensive, high priced, and in some instances inconclusive. With increasing demand of healthcare systems and scarcity of resources, there is a need for automated, reliable, non-invasive diagnostic support systems [1].

Machine learning (ML) algorithms have gained substantial attention in the medical field in recent years as they can discover patterns in large and complex datasets. One such algorithms can be trained to learn the features of a patients [excellent/poor] condition – say characteristics like age, sex, type of chest pain, cholesterol level, blood pressure, and predict whether a patient exhibits heart disease or not. ML models, by learning from historical data, make rapid and accurate predictions, aiding clinicians in making better and more consistent decisions [2].

Many machine learning ML techniques have been used to heart disease classification task such as Logistic regression (LR), decision tree (DT), Support Vector Machine (SVM),K-Nearest Neighbors(KNN), RANDOM FOREST (RF), and ensemble methods. Though these models can achieve decent performance, they could suffer from the inability to capture nonlinear feature interaction or high dimensional feature engineering. On the other hand, neural networks deep learning models in particular have the capacity to automatically learn complex input-output mappings and have outperformed the state of the art in medical prediction problems [3-7].

In this paper, Normalized Clinical Feature Neural Net (NCF-NN) _Net, a neural network-based heart disease classification model, has been proposed. NCF-NN is a feed-forward neural network having two layers of 10 neurons each with ReLU (Rectified Linear Unit) activation function [8] [9] [10].This is important to assemble a structured dataset of patient medical records, on which the model is trained without dimensionality reduction or feature exclusion, preserving the clinical completeness and interpretability. The data standardization, which was applied before training in order to normalize the feature distributions to improve the stability of the learning process.

II. Dataset Description and Feature Extraction

The dataset used in this study consists of clinical records of people who underwent examination for possible cardiovascular diseases. The dataset consists of the array of demographic, physiological, and diagnostic measurements with each instance representing a single patient serving as

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 5, pp 65-70, May-2025

https://doi.org/10.47001/IRJIET/2025.905008

predictor variables. The selected features are clinical variables which are related to heart disorders and are expected to help in further prediction. This dataset contains one binary target variable and 13 predictors. The response variable cardiovascular condition (binary): Presence (1) or absence (0). Table 1 contains descriptions of the predictor variables. These characteristics were selected because they are available in the majority of cardiological assessments and have shown a known relevance in comparable clinical trials [11].

Table 1: Description of Extracted Features from Clinical Dataset

Feature Name	Description
Age	Age of the patient in years. This feature helps assess age-related risk factors.
Sex	Gender of the patient (1 = male, 0 = female). Cardiovascular risk often differs by gender.
Chest Pain Type (cp)	Represents four types of chest pain, with higher values generally indicating more severe symptoms.
Resting Blood Pressure (trestbps)	Systolic blood pressure in mm Hg measured at rest. Elevated levels are a key risk factor.
Serum Cholesterol (chol)	Cholesterol level in mg/dL. High values are associated with atherosclerosis and heart disease.
Fasting Blood Sugar (fbs)	Indicates whether fasting blood sugar is >120 mg/dL (1 = true, 0 = false). Used to screen for diabetes.
Resting ECG Results (restecg)	Results from the resting electrocardiogram, coded into three levels. Reflects possible heart abnormalities.
Max Heart Rate Achieved (thalach)	Maximum heart rate achieved during exercise. Lower values may indicate underlying heart problems.
Exercise-Induced Angina (exang)	Binary variable (1 = yes, 0 = no) indicating whether angina occurs during physical activity.
ST Depression (oldpeak)	Degree of ST depression induced by exercise relative to rest. A diagnostic marker for ischemia.
Slope of ST Segment (slope)	Describes the slope of the ST segment during peak exercise (upsloping, flat, or downsloping).
Number of Major Vessels (ca)	Count of major vessels (0–3) colored by fluoroscopy. Higher counts are associated with increased risk.
Thalassemia (thal)	Thalassemia test result (normal, fixed defect, reversible defect). Indicates oxygen transport abnormalities.

Together these features provide a robust characterization of each subject's clinical profile and enable a granular analysis of their cardiovascular health. No derived or synthetic features were created, and all features were retained in native form as present in the clinical records during feature extraction.

The output of this pipeline is a structured dataset that can be used as input into classification models due to the extraction of consistent and domain-specific features that can inform the prediction of the target condition. The features vary in types and granularity, which make the predictive framework proposed in this work become more robust.

III. Network Architecture Design

NCF-NN is designed with architecture to tackle the task of binary classification, where the goal is identifying the presence or absence of CVDs from clinical features. The network is designed to be both accurate and computationally efficient, enabling embedding into medical decision-support systems.

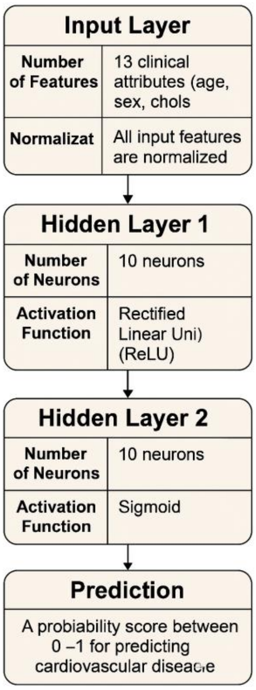


Figure 1: NCF-NN flow chart

3.1 Input Layer

NCF-NN takes thirteen clinical attributes as features in its input layer. Such as age and sex, resting blood pressure, cholesterol level, fasting blood sugar, ECG result, and thalassemia test results, etc. Because we want to ensure that all features contribute proportional during training we normalize every input variable before feeding the model.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 5, pp 65-70, May-2025

https://doi.org/10.47001/IRJIET/2025.905008

3.2 Hidden Layers

The Two fully connected hidden layers. There are ten neurons in each layer which turns out to be just about the right balance between model complexity and learning. It acts as a neuron which applies Rectified Linear Unit (ReLU) as its activation function. The non-linearity in deep learning allows the model to learn complex and non-linear relationships in the clinical data quantitatively, and the rectified linear unit (ReLU) function is usually employed as an activation function to implement layer-wise feature mapping for deep learning. By using ReLU, we also avoided the vanishing gradient problem and it converges faster during training [10].

3.3 Output Layer

The output layer is a single neuron with sigmoid activation function. The model returns a probability score of between 0 and 1 which estimates how likely it is a patient has a cardiovascular disease. Have been used a threshold equal to 0.5 to classify these predictions: equal to or higher will be classified as positive for heart disease, while below it will be classified as negative [12] [13].

3.4 Regularization and optimization

Have been used L2 regularization (ridge penalty) during training to avoid overfitting and improve generalization. This helps prevents the model from putting too much weight on one feature. A binary cross-entropy loss function ((2)) is used, which is appropriate for this type of binary classification task, and a stochastic optimization method is implemented to train the network. It works by limiting the training process to a predefined number of iterations that guarantees that the model converges, but without overfitting to the training data [14] [15] [16].

3.5 NCF-NN parameters

The design has been guided by a combination of ReLU and sigmoid activations in the architecture of NCF-NN, which allows the model to learn complex interactions while providing stable and probabilistic outputs. In addition, the use of regularization and normalization also ensures a reliable model that can generalize on unseen data well.

Table 2: structured of NCF-NN parameters and respective descriptions

Component	Parameter	Description
Input Layer	Number of Features	13 clinical attributes (age, sex, blood pressure, cholesterol, blood sugar, ECG outcomes, thalassemia)
	Normalization	All input features are normalized to a common scale for stable and efficient training.
Hidden Layer 1	Number of Neurons	10 neurons
	Activation Function	Rectified Linear Unit (ReLU)
	Purpose	To learn non-linear relationships and interactions among the input features.
Hidden Layer 2	Number of Neurons	10 neurons
	Activation Function	Rectified Linear Unit (ReLU)
	Purpose	To refine the learned features and enhance the model's ability to generalize.
Output Layer	Number of Neurons	1 neuron
	Activation Function	Sigmoid activation function
	Purpose	Outputs a probability score between 0 and 1 for predicting cardiovascular disease.
	Threshold	0.5, where values ≥ 0.5 predict positive, and values < 0.5 predict negative for cardiovascular disease.
Regularization	Method	L2 Regularization (Ridge Penalty)
	Purpose	Prevents overfitting by penalizing large weights during training.
Optimization	Algorithm	Stochastic Gradient Descent (SGD) with binary cross-entropy loss
	Purpose	Minimizes the binary cross-entropy loss to improve model prediction accuracy.
Training Constraints	Max Iterations	Fixed number of iterations for training to avoid overfitting and unnecessary computation.

International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 9, Issue 5, pp 65-70, May-2025

https://doi.org/10.47001/IRJIET/2025.905008

This table gives a clear overview of each component in NCF-NN, describing the associated parameters and their roles in the network architecture. This study very important in health care and new technology [15-23]

IV. Discussion of NCF-NN Performance Evaluation

Cardio_Net, a bilayer neural network for heart disease prediction, are presented as numerical measures and visualizations. Figure 2, the confusion matrix shows the model correctly classified 489 of the 499 non-heart disease cases and 516 of the 526 heart disease cases correctly giving it a surprisingly low number of misclassifications, only 10 false positives and 10 false negatives as show in figure 2. This shows the model has a fair trade off between sensitivity and specificity which is able to minimize class 1 error (false positive) and class2 error (false negative).

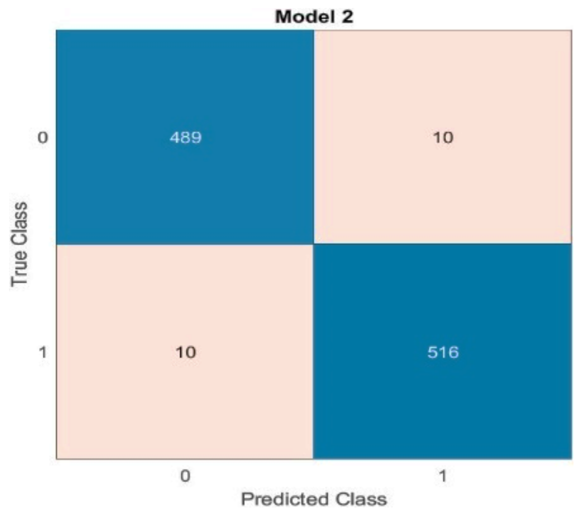
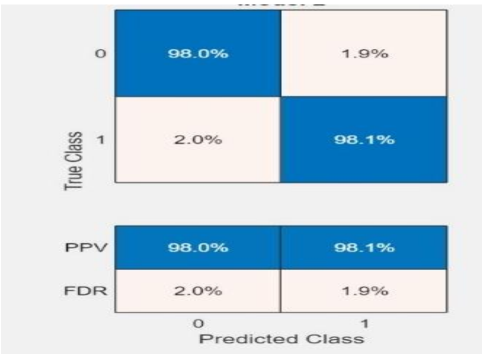


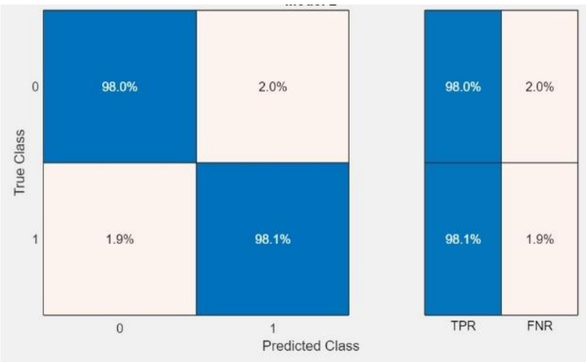
Figure 2: shows the confusion matrix for model correctly and non-correct prediction

The class 1 (for presence of heart disease) precision (Positive Predictive Value) is 98.1% which indicates that Cardio _Net is correct 98.1% of the time when it predicts a patient with heart disease. Similarly for class 0 (no heart disease), the net have precision being 98.0%. These values are important since it means that the model is more reliable in terms of positive or negative predictions.

This is further supported by the very low False Discovery Rate (FDR) at 2.0% for class 0 and 1.9% for class 1, indicating that among all systematic cases predicted to be positive, only a very small proportion are in fact false. These low FDR indicate a high confidence in predictions and are critical when applied to the healthcare setting to minimize unnecessary interventions or missed treatment as show in figure 3 (a-b).



(a)



(b)

Figure 3: Illustrate the PPV, FDR, TPR and FVR

The model strength is also corroborated by the Receiver Operating Characteristic (ROC) curve shown in figure 4. The ROC curve shoots up to the top-left quadrant, where the TPR is very high and the FPR is very low. The AUC is 0.99 which indicates almost perfect separability between heart disease patients and heart disease absent patients. The classifier points at (0.02, 0.98) denotes a false positive rate of only 2% and true positive rate of 98% meaning only 2% of data points not indicating disease are misclassified as having disease (FP) while 98% of data points indicating disease are flagged correctly TP.

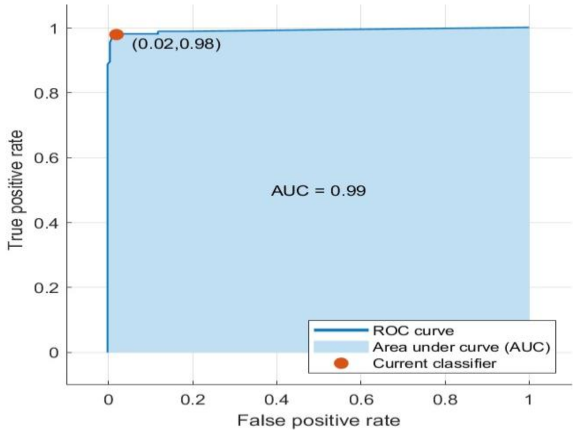


Figure 4: The ROC curve evaluation for heart disease model

ISSN (online): 2581-3048
Volume 9, Issue 5, pp 65-70, May-2025
https://doi.org/10.47001/IRJIET/2025.905008

An average accuracy of 98.05% confirms that these metrics are consistent and the model scores similarly across both the classes. Note that all these metrics make sense in synergy and in quintessence indicate that the model is neither biased to one class nor overfitting.

This impressive performance is in large part due to the design of the NCF-NN — maintaining the interpretability of medical variables by taking all clinical features as is without PCA and also standardized inputs. Besides being speedy in training (having 6.15 seconds per training) and computationally efficient (having accuracy with 96.25 accuracy within fraction of second), NCF-NN is also very decent for real-life applications in the clinical decision support systems.

V. Conclusion

In conclusion, the proposed NCF-NN model has effectively leveraged a straightforward nonetheless potent neural network architecture to anticipate the existence of heart disease with high accuracy and dependability. By incorporating all clinically pertinent features and applying standardization and regularization strategies, the model attained superb performance across multiple assessment metrics, such as precision, recall, and area under the curve. The findings underscore the potential of NCF-NN high performance in predictive diagnosis, particularly when utilized to structured medical records.

REFERENCES

[1] Ahmad AA, Polat H. “Prediction of Heart Disease Based on Machine Learning Using Jellyfish Optimization Algorithm”. *Diagnostics (Basel).* 2023 Jul 17; 13(14): 2392. doi: 10.3390/diagnostics13142392. PMID: 37510136; PMCID: PMC10378171.

[2] Qassim, H. Early Prediction of Stroke Risk Using Machine Learning Approaches and Imbalanced Data. *NTU Journal of Engineering and Technology*, 4(1), 2025.

[3] Al-Nima, R. R. O., Al-Hatab, M. M. M., & Qasim, M. A. (2023). An artificial intelligence approach for verifying persons by employing the deoxyribonucleic acid (DNA) nucleotides. *Journal of Electrical and Computer Engineering*, 2023(1), 6678837.

[4] Hameed, D. M., Aziz, R. R., Malla, M. A., Al-Hatab, M. M. M., & AL-Nima, R. R. O. (2025). Performance Evaluation of SVM Classifiers for Atrial Fibrillation Detection. *Iraqi Journal of Science*.

[5] Al-Hatab, M. M. M., Alhashim, M. A., Fadhil, M. A., Hasan, A. J. A. R., & Al-Sultan, T. G. (2022). Innovative Non-Invasive Blood Sugar Level Monitoring for Diabetes Using UWB Sensor. *Journal of Optoelectronics Laser*, 41(4), 422-437.

[6] Al-Hatab, M. M. M., &AlNima, M. Z. S. (2023). Hematological classification of white blood cells by exploiting digital microscopic images. *Eur Res Bull,* 18, 44-52.

[7] Veisi, Hadi, Hamid Reza Ghaedsharaf, and Morteza Ebrahimi. "Improving the performance of machine learning algorithms for heart disease diagnosis by optimizing data and features." *Soft computing journal* 8, no. 1 (2021): 70-85.

[8] Parisi, L., Neagu, D., Ma, R., & Campean, F. (2022). Quantum ReLU activation for convolutional neural networks to improve diagnosis of Parkinson’s disease and COVID-19. *Expert systems with applications*, 187, 115892.

[9] Dutta, A., Batabyal, T., Basu, M., & Acton, S. T. (2020). An efficient convolutional neural network for coronary heart disease prediction. *Expert Systems with Applications*, 159, 113408.

[10] Lenin, S. T., and K. Venkatasalam. "Effective Classification of Heart Disease Using Convolutional Neural Networks." *Circuits, Systems, and Signal Processing* 44, no. 2 (2025): 911-935.

[11] https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset

[12] Ajam, Noura. "Heart diseases diagnoses using artificial neural network." *IISTE Network and Complex Systems* 5, no. 4 (2015).

[13] Hasan, Tabreer T., Manal H. Jasim, and Ivan A. Hashim. "Heart disease diagnosis system based on multi-layer perceptron neural network and support vector machine." *International Journal of Current Engineering and Technology* 77, no. 55 (2017): 2277-4106.

[14] Patro, Sibo Prasad, Gouri Sankar Nayak, and Neelamadhab Padhy. "Heart disease prediction by using novel optimization algorithm: A supervised learning prospective." *Informatics in Medicine Unlocked* 26 (2021): 100696.

[15] Parikh, Vanssh, Bhoomi Sharma, Arnav Byotra, and Amarjit Malhotra. "Optimizing Heart Disease Prediction Using a Hybrid Dynamic Swarm Evolution Approach." *SN Computer Science* 5, no. 8 (2024): 1104.

[16] Mondal, Subhash, Ranjan Maity, and Amitava Nag. "An efficient artificial neural network-based optimization techniques for the early prediction of coronary heart disease: comprehensive analysis." *Scientific Reports* 15, no. 1 (2025): 4827.

[17] Al-Hatab, M. M. M., Al-Obaidi, A. S. I., & Al-Hashim, M. A. (2024). Exploring CIE lab color characteristics

ISSN (online): 2581-3048
Volume 9, Issue 5, pp 65-70, May-2025
https://doi.org/10.47001/IRJIET/2025.905008

for skin lesion images detection: a novel image analysis methodology incorporating color-based segmentation and luminosity analysis. *Fusion: Practice and Applications*, 15(1), 88-97.

[18] Fathel, W. R., Al-Obaidi, A. S. I., Qasim, M. A., & Al-Hatab, M. M. (2023). Skin cancer detection using K-means clustering-based color segmentation. *Texas J Eng Technol*, 18(2770-4491), 46-52.

[19] Alhelal, Dheyaa, Ahmed Khazal Younis, and Ruaa H. Ali Al-Mallah. "Detection of brain stroke in the MRI image using FPGA." *TELKOMNIKA (Telecommunication Computing Electronics and Control)* 19.4 (2021): 1307-1315.

[20] Al-Mallah, Ruaa H. Ali, DheyaaAlhelal, and Razan Abdulhammed. "ASSAS: An automatic smart students attendance system based on normalized cross-correlation." *Bulletin of Electrical Engineering and Informatics* 10.2 (2021): 732-741.

[21] Daoud, Raid W., Ruaa H. Ali Al-Mallah, Mahmood Sh Majeed, and Yaareb MBI Al-khashab. "Design and Simulate an Attenuator for Multi Types Optical Fiber Using Neural Networks." *International Journal of Enhanced Research in Science, Technology & Engineering* 8, no. 5 (2019): 19-26.

[22] Al-Jabbar, A., Entisar, Y., Mohamedsheet Al-Hatab, M. M., Qasim, M. A., Fathel, W. R., & Fadhil, M. A. (2023). Clinical Fusion for Real-Time Complex QRS Pattern Detection in Wearable ECG Using the Pan-Tompkins Algorithm. *Fusion: Practice & Applications*, 12(2).

[23] Alrawi, R. M. S., & Basheer, N. M. (2025). Pediatric Radiology: An Analysis of AI-Powered Bone Age Determination Methods. *NTU Journal of Engineering and Technology*, 4(1).

Citation of this Article:

Ruaa H. Ali Al-Mallah, & Marwa Mawfaq Mohamedsheet Al-Hatab. (2025). Normalized Clinical Feature Neural Net (NCF-NN) for Cardiovascular Prediction. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 9(5), 65-70. Article DOI https://doi.org/10.47001/IRJIET/2025.905008

****