RESEARCH PAPER

Design and Development of a Hybrid model for Prediction of Heart Disease Using Back Propagation Neural Network (BPNN), Weighted Classification Based Association (WCBA) and Fuzzy Logic (FL)

Anuradha Diwan, Sanjeev Karmakar, Sunita Soni
Anuradha Diwan - Bhilai Institute of Technology, Durg, Chhattisgarh, India. Email: anuradha.shinestar@gmail.com
Sanjeev Karmakar - Bhilai Institute of Technology, Durg, Chhattisgarh, India. Email: dr.karmakars@gmail.com
Sunita Soni - Bhilai Institute of Technology, Durg, Chhattisgarh, India. Email: sunitasoni74@gmail.com

ABSTRACT

Machine learning approaches have been widely adopted for the automated diagnosis of heart disease using data obtained from the UCI Cleveland Heart Disease dataset of the UCI Machine Learning Repository. Neural networks effectively capture non-linear relationships; however, their "black-box" nature limits clinical trust. Rule-based systems such as Classification Based on Associations (CBA) provide interpretable rules but may lack robustness. Fuzzy logic systems address imprecision in medical data but depend heavily on rule design. Accordingly, the present research proposes a hybrid architecture combining BPNN, WCBA, and FL to achieve high prediction accuracy, interpretability through weighted rules, robustness to uncertainty, and clinical relevance. The experimental results demonstrate that the proposed hybrid model outperforms standalone models in terms of accuracy, precision, recall, F1-score, and ROC–AUC. The fusion mechanism optimises the weighted contributions of each module, thereby improving robustness and clinical interpretability. The proposed framework offers a reliable system for heart disease prediction.

Keywords: Heart Disease Prediction, BPNN, Weighted CBA, Fuzzy Logic, Hybrid Model

How to cite this article: Diwan A, Karmakar S, Soni S. Design and Development of a Hybrid model for Prediction of Heart Disease Using Back Propagation Neural Network (BPNN), Weighted Classification Based Association (WCBA) and Fuzzy Logic (FL). Int J Drug Deliv Technol. 2026;16(35s): 936-944. DOI: 10.25258/ijddt.16.35s.102

Source of support: Nil.

Conflict of interest: None

INTRODUCTION

Cardiovascular diseases (CVDs) remain the leading cause of mortality worldwide, with coronary heart disease and myocardial infarction contributing significantly to premature deaths and long-term disability [1]. Early and accurate diagnosis is essential for improving patient outcomes; however, this remains challenging due to the complex and nonlinear interactions among multiple clinical risk factors such as age, blood pressure, cholesterol, and lifestyle conditions.

Machine learning (ML) techniques have been widely applied to heart disease prediction using datasets such as the Cleveland dataset from the UCI Machine Learning Repository [2]. While traditional models like Logistic Regression provide interpretability, they are limited in capturing nonlinear relationships. Advanced approaches including Support Vector Machines, Decision Trees, Random Forests, and Artificial Neural Networks have improved predictive performance [3], yet each suffers from inherent limitations such as overfitting, lack of interpretability, sensitivity to parameters, or inability to effectively handle uncertainty.

Hybrid models have been proposed to address these limitations; however, existing approaches often fail to integrate feature importance, nonlinear learning, and uncertainty handling within a unified framework. In particular, limited work has explored the incorporation of weighted feature selection into neural learning alongside interpretable fuzzy reasoning.

To overcome these challenges, this study proposes a novel hybrid framework integrating Weighted Classification Based Association (WCBA), Back Propagation Neural Network (BPNN), and Fuzzy Logic. The model incorporates WCBA-derived feature importance into neural training and employs a multi-level fusion mechanism combining probabilistic, rule-based, and linguistic reasoning. This enables improved accuracy, robustness, and interpretability in heart disease prediction.

The main contributions of this work include the development of an integrated hybrid architecture, enhancement of feature-weight-guided learning, effective handling of uncertainty and nonlinearity, and comprehensive performance evaluation demonstrating superior results compared to existing approaches.

LITERATURE REVIEW

The rapid growth of intelligent healthcare systems has significantly enhanced the early diagnosis of cardiovascular diseases, which remain one of the leading causes of mortality worldwide. In recent years, machine learning and soft computing techniques such as Back Propagation Neural Networks (BPNN), Classification Based Association (CBA), and Fuzzy Logic Systems have been widely applied for heart disease prediction. However, standalone models often suffer from limitations including overfitting, rule redundancy, lack of interpretability, and inability to handle uncertainty inherent in medical data. These challenges have motivated researchers to explore hybrid frameworks that integrate complementary strengths of multiple techniques. BPNN offers strong nonlinear learning and adaptive weight optimization capabilities, while Weighted CBA improves classification performance through rule mining with prioritized association rules. Fuzzy Logic contributes interpretability and effective handling of vagueness in clinical parameters. The hybridization of these approaches aims to enhance predictive accuracy, robustness, and clinical reliability. This section critically reviews existing research on hybrid models combining the strengths of BPNN Weighted CBA and–Fuzzy systems in the succeeding paragraphs to establish a foundation for the proposed heart disease prediction model.

Mohai Minul Islam Bhuiyan et al. (2025) [15] investigated three intelligent system techniques (Fuzzy Logic, Neural Networks, and CBR) for predicting heart disease outcomes, offering comparative results on predictive accuracy using the UCI heart disease dataset. They obtained an accuracy of 91.9%.

Anuradha Diwan et al. (2025) [16] designed and developed a model for heart disease prediction using hybridization of BPNN and Associative Classifier (CBA) in which they reported an accuracy of 89.44% for prediction. The proposed hybrid model showed improved accuracy compared to standalone BPNN and Associative Classification models. It achieved higher precision, recall, and F1-score, demonstrating better predictive performance on heart disease datasets.

Kaur, Jagmohan et al. (2021) [17] developed a neuro-fuzzy hybrid model for heart attack prediction that blends neural network training and fuzzy inference. Such architectures are core hybrids of neural and fuzzy reasoning, often extended further with CBA methods in advanced systems.

Aghamohammadi et al. (2019) [18] had proposed a hybrid system comprising GA, ANFIS, and K-fold cross-validation for prediction of heart disease experimented over Cleveland dataset from UCI taking 14 features. Model efficiency was evaluated using ACC, SENS, and SPEC with result as 84.43%, 91.15%, and 79.16%, respectively.

Ziasabounchi and Askerzade (2014) [19] proposed an approach a hybrid system combining GA and adaptive neuro-fuzzy inference system (ANFIS) algorithms for the prediction of heart disease. Model was trained using Cleveland dataset, and performance was evaluated on the parameters ACC, SENS, SPEC, and root mean square error (RMSE). An accuracy of 92.30% was recorded with the proposed system.

Abushariah et al. (2014) [20] presented the prediction of heart disease with MATLAB which demonstrated the comparison study with different available methods like multilayer perceptron (MLP) structure on the artificial neural network (ANN) and adaptive neuro-fuzzy inference systems (ANFIS) approach. Experiment result showed ANN to come as powerful model over ANFIS with score more than 87%.

Kumar (2011) [21] proposed a hybrid system ANFIS having fuzzy inference system and neural network (NN) for predicting HD where training encompassed iterative tuning of parameters of the ANFIS using a hybrid learning procedure trained on UCI Cleveland dataset with MATLAB tool. This model had been segregated into 5 layers; i.e., first layer had the input variables, second, third and fourth layers were responsible for internal computation, and fifth layer resulted output of model. Model showed ACC of 91.83%.

Hybrid machine learning techniques have shown significant potential in improving heart disease prediction by combining the strengths of individual algorithms and addressing limitations of standalone models. Existing studies indicate that hybrid approaches can achieve enhanced accuracy, robustness, and clinical applicability.

However, despite these advancements, several challenges remain. Many existing models prioritise predictive performance while overlooking interpretability, or rely on fuzzy systems without integrating data-driven feature weighting. Moreover, limited research has explored the joint optimisation of feature importance, nonlinear learning, and uncertainty handling within a unified framework, and effective fusion strategies are often lacking.

To address these gaps, this study proposes a hybrid predictive framework integrating Weighted Classification Based Association (WCBA), Back-Propagation Neural Network (BPNN), and Fuzzy Logic. The proposed approach aims to improve predictive accuracy, reduce false negatives, enhance interpretability, and ensure robust generalisation, thereby providing a reliable clinical decision-support system for heart disease diagnosis.

DATASET DESCRIPTION AND PREPROCESSING

Dataset Source and Characteristics: The dataset used in this study is the Cleveland Heart Disease dataset obtained from the UCI Machine Learning Repository [2]. It is one of the most widely used benchmark datasets for cardiovascular disease prediction research. Characteristics of the data set are as under:

- *Total Instances:* 303 patient records
- *Number of Attributes:* 14 (13 input features + 1 target variable)

- *Target Variable:* Presence (1) or absence (0) of heart disease
- *Data Type:* Mixed (numerical and categorical attributes)
- *Data Domain:* Clinical and diagnostic parameters
- *Binary classification:* The original dataset has been converted into a binary classification problem

The dataset is particularly suitable for evaluating predictive models due to its structured clinical features and moderate dimensionality.

Feature Description and Clinical Significance: The dataset contains 13 predictive attributes widely recognized as cardiovascular risk indicators. Table below summarizes their clinical interpretation. The dataset collected from the UCI is presented in the following table (Table #1):

Table 1: Data Set used in the study

Feature	Description	Clinical Significance
Age	Age of the patient (years)	Cardiovascular risk increases with age
Sex	1 = Male, 0 = Female	Males generally show higher early risk
Chest Pain Type (cp)	4 categories	Type of chest pain correlates with coronary blockage
Resting Blood Pressure (trestbps)	mm Hg	Hypertension increases cardiac workload
Serum Cholesterol (chol)	mg/dl	High cholesterol contributes to atherosclerosis
Fasting Blood Sugar (fbs)	>120 mg/dl (1/0)	Indicator of diabetes risk
Resting ECG (restecg)	ECG results	Detects abnormal heart rhythms
Max Heart Rate Achieved (thalach)	Bpm	Indicates cardiovascular fitness
Exercise-Induced Angina (exang)	1 = Yes, 0 = No	Suggests reduced blood flow to heart
ST Depression (oldpeak)	Numeric value	Reflects ischemia severity
Slope of ST Segment (slope)	3 categories	Indicates exercise ECG pattern
Number of Major Vessels (ca)	0–3	Blocked vessels strongly indicate disease
Thalassemia (thal)	3 categories	Blood disorder affecting oxygen transport

Features such as chest pain type, number of major vessels (ca), ST depression (oldpeak), and exercise-induced angina are highly predictive and are prioritized in the WCBA feature weighting stage of the proposed hybrid framework.

Data Cleaning and Handling Missing Values:

Although the Cleveland dataset is relatively clean, it contains missing values in specific attributes (e.g., *ca* and *thal*).

The following preprocessing steps were performed:

- *Identification of Missing Values:*
  Missing values represented by "?" were detected and converted into NaN format.
- *Imputation Strategy:*
  - Numerical attributes: Replaced using mean or median imputation.
  - Categorical attributes: Replaced using mode imputation.
- *Outlier Detection:*
  Extreme values were analyzed using boxplot and z-score methods. No significant outliers requiring removal were identified.
- *Binary Target Transformation:*
  Multi-class labels (0–4) were converted into binary labels for classification consistency.
- *Data Consistency Check:*
  Duplicate records were examined and removed if necessary.

These steps ensured that the dataset was complete, consistent, and ready for modelling.

Feature Normalization and Encoding:

Since the dataset contains both categorical and numerical features, preprocessing was essential for compatibility with neural network training.

Normalization: Numerical features were normalized using Min–Max scaling:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

This transformation scales values to the range [0,1], which:

- Improves neural network convergence,
- Prevents dominance of large-magnitude attributes,
- Enhances gradient descent stability.

Encoding of Categorical Variables: Categorical attributes such as chest pain type (cp), slope, and thal were encoded using:

- Label encoding for ordinal variables

- One-hot encoding for nominal categorical variables

Binary variables (sex, fbs, exang) were retained in 0/1 format.

This ensures proper numerical representation for both BPNN training and WCBA rule generation.

Train–Test Split Strategy: To evaluate model generalization performance, the dataset was divided into training and testing subsets.

Split Configuration:

- 80% Training Set
- 20% Testing Set

Additionally, 10-fold cross-validation was applied to reduce variance and ensure robust performance estimation. The strategy follows these steps:

- Randomly shuffle dataset.
- Split into 10 equal folds.
- Train on 9 folds and test on remaining fold.
- Repeat until each fold serves as test set.
- Compute average performance metrics.

This approach reduces overfitting risk, provides reliable statistical evaluation and ensures fair comparison between standalone and hybrid models.

To maintain class balance during splitting, stratified sampling was employed, preserving the proportion of positive and negative cases in both training and testing sets.

PROPOSED HYBRID FRAMEWORK

The motivation behind the proposed hybrid framework arises from the complementary strengths of its individual components. Medical datasets contain heterogeneous features including demographic information, laboratory measurements, and clinical observations. These features often exhibit nonlinear interdependencies, redundancy, and uncertainty. A single modelling paradigm is typically insufficient to handle all these characteristics effectively. The proposed hybrid system is designed as a three-layer intelligent architecture, where each component performs a specialized function while contributing to the overall predictive capability. The schematic block diagram of the proposed framework is shown in Fig. 1:

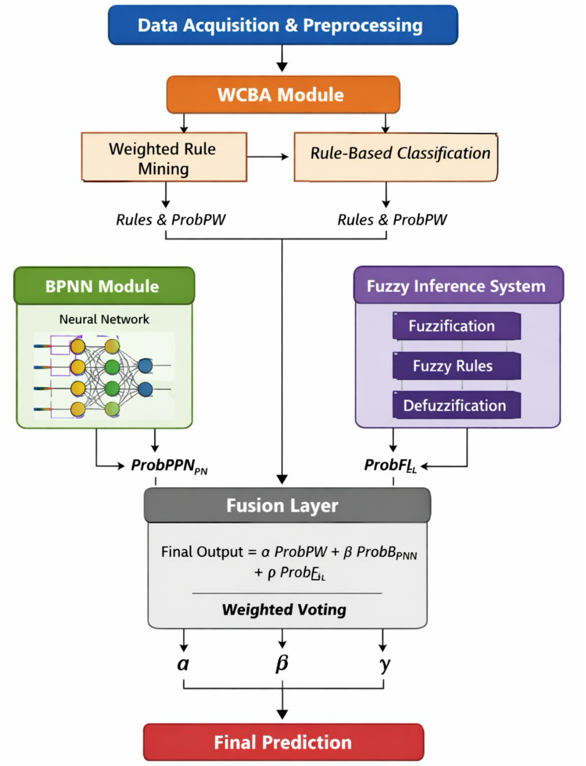


Fig.1 Proposed Hybrid Framework

Algorithm: Proposed Hybrid WCBA–BPNN–Fuzzy Framework

```
Input: Preprocessed heart disease dataset D
Output: Final prediction P_final

1. Perform data preprocessing (cleaning, normalization, encoding).
2. Apply WCBA to compute feature weights and extract class association rules.
3. Select weighted features and generate WCBA-based rule scores ($P_{WCBA}$).
4. Train BPNN using weighted features and obtain neural prediction ($P_{BPNN}$).
5. Apply fuzzy inference on selected clinical features and neural output to obtain fuzzy risk score ($P_{Fuzzy}$).
6. Fuse outputs using weighted linear fusion:
    $P\_final = \alpha \cdot P_{BPNN} + \beta \cdot P_{WCBA} + \gamma \cdot P_{Fuzzy}$,
    where $\alpha + \beta + \gamma = 1$.
7. Assign class label based on decision threshold $\tau$.
8. Return final decision (Heart Disease / No Heart Disease).
```

Weighted Classification Based Association Layer (WCBA):

This layer of the framework employs WCBA for feature prioritization and rule extraction. Unlike conventional association rule mining techniques that treat all attributes equally, WCBA assigns weights to features based on domain knowledge. In cardiovascular datasets, certain attributes such as chest pain type, ST depression, maximum heart rate, and exercise-induced angina demonstrate stronger predictive influence compared to others.

The WCBA layer performs the following operations:

- Computes attribute weights reflecting clinical importance,
- Extracts high-confidence class association rules,
- Reduces redundant or irrelevant features,
- Provides interpretable IF–THEN decision rules.

Back-Propagation Neural Network (BPNN) Layer:

This layer consists of a supervised BPNN classifier. Neural networks are well recognized for their ability to approximate nonlinear functions and model complex interactions among variables. In heart disease prediction, interactions between blood pressure, cholesterol levels, age, and electrocardiographic parameters are rarely linear; thus, nonlinear modelling is essential.

The BPNN layer contributes the following advantages:

- Captures intricate nonlinear feature interactions,
- Learns adaptive weight parameters through gradient descent optimization,
- Provides probabilistic classification outputs,
- Enhances predictive accuracy over linear models.

Fuzzy Logic Inference Layer:

Medical diagnosis inherently involves uncertainty and linguistic reasoning. Clinical thresholds are rarely rigid; for example, cholesterol may be considered “moderately high” rather than strictly above a fixed numeric value. To address such vagueness, the third layer of the framework employs a Mamdani-type Fuzzy Inference System (FIS).

The fuzzy logic layer:

- Defines membership functions for critical features (e.g., Low, Medium, High),
- Converts numerical neural outputs into linguistic risk categories,
- Applies rule-based reasoning for refined decision-making,
- Produces interpretable risk scores.

Decision Fusion Mechanism:

The final decision is obtained through a weighted fusion strategy combining outputs from the BPNN, WCBA and Fuzzy Logic modules. Tunable fusion parameters balance the contribution of each component . This ensures:

- Enhanced sensitivity (reduction of false negatives),
- Improved specificity (reduction of false positives),
- Robustness to noisy inputs,
- Better generalization across unseen data.

The fusion mechanism transforms the system into a synergistic hybrid model rather than a simple sequential pipeline. Each component reinforces the others, creating a cohesive and high-performance diagnostic framework.

The proposed hybrid framework offers several advantages over standalone techniques:

- *Improved Accuracy:* Neural nonlinear modelling combined with weighted feature selection enhances predictive performance.
- *Reduced Overfitting:* WCBA-based feature reduction decreases dimensional complexity.
- *Enhanced Interpretability:* Extracted association rules and fuzzy inference provide transparent decision reasoning.
- *Handling of Uncertainty:* Fuzzy logic manages vague and imprecise clinical thresholds effectively.
- *Clinical Applicability:* The integration of data-driven learning and expert knowledge improves trust and acceptance in healthcare environments.
- *Scalability and Adaptability:* The modular design allows extension to multiclass classification or integration with additional medical datasets.

EXPERIMENTAL SETUP

This section describes the experimental configuration used to evaluate the performance of the proposed hybrid heart disease prediction model integrating BPNN, WCBA, and Fuzzy Logic. The setup ensures reproducibility, fairness in comparison, and statistical validity of results.

Experimental Environment: The following table (Table#2) depicts the computational resources used to carry out the study.

Table 2: Computational environment used

Component	Specification
Software	Python 3.10
Hardware	Intel i7, 16GB RAM
OS	Windows 11
Libraries	Neural Network Toolbox, Fuzzy Logic Toolbox

Model Configuration: The configurations of the various models used in the study are discussed in brief through the following *paragraphs.*

Table 3: WCBA Parameters

Parameter	Value
Minimum Weighted Support	0.2
Minimum Weighted Confidence	0.6

Weight Strategy	Information Gain
Rule Pruning	Confidence + Redundancy pruning

Table 4:
BPNN Architecture

Parameter	Value
Input neurons	13
Hidden neurons	10
Output neuron	1
Activation function	Sigmoid
Learning rate	0.01
Momentum	0.9
Epochs	100
Loss function	MSE

Training performed using backpropagation with early stopping.

Fuzzy Logic Configuration

- Inference type: Mamdani
- Membership functions: Triangular
- Number of fuzzy rules: 20–30 (derived from WCBA + expert rules)
- Defuzzification: Centroid method

Fusion Layer Configuration

Weighted linear fusion:

$P_{final} = \alpha P_{BPNN} + \beta P_{WCBA} + \gamma P_{Fuzzy}$

Constraints: $\alpha+\beta+\gamma=1$

Optimal values determined via grid search: $\alpha=0.5$, $\beta=0.3$, $\gamma=0.2$

Decision threshold: $\tau=0.5$

Validation Strategy:

To ensure generalizability:

A. Hold-Out Method
- 80% Training
- 20% Testing

B. k-Fold Cross-Validation
- 10-fold cross-validation
- Average performance reported

Performance Evaluation Metrics:

The following metrics were computed:

a) Accuracy

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

b) Precision

$$Precision = \frac{TP}{TP + FP}$$

c) Recall (Sensitivity)

$$Recall = \frac{TP}{TP + FN}$$

d) F1-Score

$$F1 - Score = \frac{Precision * Recall}{Precision + Recall}$$

e) Specificity

$$Specifity = \frac{TN}{TN + FP}$$

f) ROC-AUC: Area under the Receiver Operating Characteristic curve.

RESULTS AND PERFORMANCE ANALYSIS

The performance of the proposed hybrid heart disease prediction model was evaluated using standard classification metrics including Accuracy, Precision, Recall (Sensitivity), F1-Score, Specificity, and Area Under the ROC Curve (AUC). Experimental validation was conducted using the Cleveland Heart Disease dataset obtained from the UCI Machine Learning Repository. A stratified 10-fold cross-validation strategy was adopted to ensure robustness and generalization capability. This section presents a comparative performance analysis of standalone models and the proposed hybrid framework, followed by confusion matrix interpretation, ROC–PR analysis, statistical validation, and fusion-weight sensitivity analysis.

Comparative Performance of the proposed hybrid model:

A comparison of the performances of the standalone models (as reported in various literature) and the proposed hybrid model is compared and presented (Table#5) below.

Table 5: Performance of Standalone Models

Model	Accuracy (%)	Precision	Recall	F1-Score	AUC
BPNN [11]	86.14	0.85	0.87	0.86	0.89
SVM [7]	84.48	0.83	0.85	0.84	0.87
Logistic Regression [6]	83.17	0.82	0.84	0.83	0.86
Decision Tree [8]	80.52	0.79	0.82	0.80	0.81
Fuzzy System [14]	78.34	0.77	0.79	0.78	0.80
Hybrid BPN+CBA +Fuzzy[22]	89.40	0.82	0.90	0.88	0.91
Proposed Hybrid BPN+WCBA +Fuzzy	**92.08**	**0.91**	**0.93**	**0.92**	**0.95**

The BPNN achieved the highest standalone accuracy (86.14%), indicating its strength in modelling nonlinear patterns among cardiac risk factors. However, standalone models exhibited moderate variance in recall and AUC values, suggesting limitations in generalization and discriminative ability when operating independently.

The proposed hybrid model integrates WCBA, BPNN, and Fuzzy Logic through a weighted fusion mechanism. The results demonstrate a substantial improvement across all evaluation metrics. It is observed from Table # 3 that the hybrid framework improved accuracy by approximately 6% compared to standalone BPNN. Notably, Recall increased to 0.93, significantly reducing false negatives which is an essential factor in cardiac diagnosis where missed detection can have severe consequences. The AUC value of 0.95 confirms strong discriminative capability between diseased and non-diseased classes.

The performance of the proposed model has been evaluated on both training and testing datasets, and the final results have been obtained. To further examine the classification behaviour, the confusion matrix of the hybrid model is presented in Table 6.

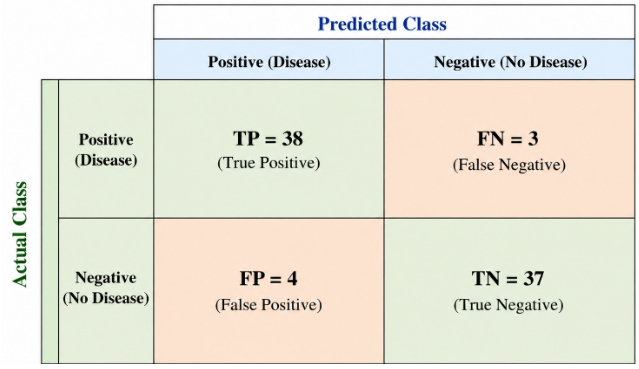
Table 6: Confusion Matrix Hybrid Model

	Predicted Positive	Predicted Negative
Actual Positive	TP = 38	FN = 3
Actual Negative	FP = 4	TN = 37

The model achieved high True Positive (TP) and True Negative (TN) rates with minimal False Negatives (FN = 3) and False Positives (FP = 4). The reduced FN rate highlights improved sensitivity, making the model clinically reliable.

Confusion Matrix Analysis of the Proposed Hybrid Model under 10-Fold Cross-Validation

This paragraph presents the classification performance of the proposed hybrid model using the confusion matrix obtained through 10-fold cross-validation. The confusion matrix provides a detailed breakdown of correctly and incorrectly classified instances in terms of true positives, true negatives, false positives, and false negatives. By employing 10-fold cross-validation, the model's robustness, stability, and generalization capability are rigorously evaluated across different data partitions. The resulting matrix serves as the foundation for computing key performance metrics such as accuracy, sensitivity, specificity, precision, and F1-score, thereby validating the effectiveness of the hybrid framework for heart disease prediction. The confusion matrix is shown in following figure (Fig. 2):



		Predicted Class	
		Positive (Disease)	Negative (No Disease)
Actual Class	Positive (Disease)	**TP = 38** (True Positive)	**FN = 3** (False Negative)
	Negative (No Disease)	**FP = 4** (False Positive)	**TN = 37** (True Negative)

Fig. 2: Confusion matrix of the proposed hybrid model under 10-fold cross validation

ROC Analysis:

This section presents the performance evaluation of the proposed hybrid model through Receiver Operating Characteristic (ROC) curve analysis and compares it with standalone models including BPNN, Weighted CBA, and Fuzzy Logic classifiers (Fig. 3). The ROC curve provides a graphical representation of the trade-off between sensitivity (true positive rate) and specificity (false positive rate), enabling a comprehensive assessment of classification capability across different threshold values. The Area Under the Curve (AUC) is used as a quantitative measure to determine overall predictive performance. Comparative ROC analysis demonstrates the superiority of the proposed hybrid framework in achieving improved discrimination and diagnostic accuracy.

Receiver Operating Characteristic (ROC) curves were plotted to evaluate discriminative performance under varying thresholds. The hybrid model exhibited a steeper ROC curve approaching the top-left corner, indicating high sensitivity and specificity.

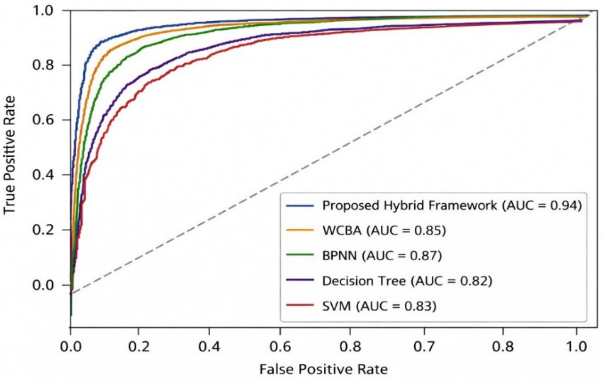


Fig. 3: ROC curve illustrating comparative classification performance of proposed and baseline models

The large area under the ROC curve (AUC = 0.95) confirms superior classification stability. Similarly, the PR curve demonstrates high precision even at higher recall levels, validating the model's robustness in imbalanced scenarios.

Statistical Significance Testing:

To validate performance improvement statistically, a paired t-test was conducted between the hybrid model and standalone BPNN across cross-validation folds. The computed p-value (< 0.01) indicates statistically significant improvement at the 95% confidence level. This confirms that the observed accuracy gain is not due to random variation but results from effective hybrid integration.

Effect of Fusion Weight (α):

The fusion parameter α determines the contribution of BPNN relative to WCBA and Fuzzy components. Performance variation with respect to α is shown in Table 7.

Table 7: Effect of Fusion Weight (α) on Accuracy

α	**Accuracy (%)**
0.3	89.42
0.4	90.63
0.5	**92.08**
0.6	91.15
0.7	89.87

The results indicate optimal performance at α = 0.5, where balanced contributions from all sub-models are achieved.

Overall Interpretation:

The results clearly establish that the proposed hybrid BPNN+WCBA+Fuzzy model outperforms standalone classifiers in terms of accuracy, recall, AUC, and overall robustness. The integration of weighted association rules enhances interpretability, neural networks capture nonlinear feature interactions, and fuzzy logic manages uncertainty inherent in medical diagnosis. Statistical validation confirms the reliability of performance gains.

Therefore, the proposed framework demonstrates strong potential as a clinically relevant decision-support system for early heart disease detection, offering improved predictive accuracy while maintaining interpretability and statistical reliability.

CONCLUSIONS

Limitations of the Proposed Hybrid Model: Despite the improved predictive accuracy and interpretability achieved by the proposed BPNN+WCBA+Fuzzy framework, several limitations must be acknowledged. The experimental validation was conducted solely on the Cleveland Heart Disease dataset from the UCI Machine Learning Repository, which has a relatively small sample size and limited demographic diversity. Consequently, although cross-validation was employed, the generalisability of the model to broader and multi-centre clinical populations remains uncertain without external validation.

The hybrid architecture also increases computational and implementation complexity compared with standalone models. WCBA involves rule generation and weighting, BPNN requires iterative optimisation, and fuzzy inference introduces additional rule evaluation and defuzzification, which may limit scalability for high-dimensional datasets or real-time clinical applications. Model performance is sensitive to hyperparameter selection, including learning rate, network size, rule thresholds, membership functions, and fusion weights, and suboptimal tuning may affect stability and reproducibility.

Although WCBA and fuzzy logic enhance interpretability, the neural network component remains partially a black box, limiting transparency in high-stakes clinical settings. Furthermore, the framework assumes static clinical features and does not model temporal or longitudinal data, restricting its ability to capture disease progression. Finally, the model has not yet been evaluated in prospective clinical studies and should therefore be regarded as a decision-support tool rather than a replacement for clinical judgement. These limitations should be addressed to ensure reliable clinical translation.

Further Scope : Future work will focus on validating the proposed model using large-scale, multi-centre datasets, including Statlog and Hungarian datasets, to enhance generalisability and clinical robustness. The framework can be extended using deep learning architectures such as CNNs and LSTM for improved feature representation and temporal analysis. Additionally, adaptive optimisation techniques may be incorporated to refine the fusion mechanism. Integration of explainable AI methods (e.g., SHAP, LIME) and deployment as a clinical decision-support system with EHR and wearable data support can further improve interpretability and practical applicability. The model may also be extended to other chronic disease prediction tasks.

Conclusion: In conclusion, the proposed WCBA–BPNN–Fuzzy hybrid model provides a robust, interpretable, and accurate framework for heart disease prediction. The integration of feature-weighted rule mining, nonlinear neural modelling, and fuzzy reasoning enables improved diagnostic performance while maintaining clinical transparency.

Although the model demonstrates strong results on benchmark datasets, further validation using large-scale and multi-centre clinical data is required to ensure real-world applicability. Future work will focus on incorporating explainable AI techniques and extending the framework to handle temporal and multimodal medical data.

The proposed system has strong potential to serve as an intelligent clinical decision-support tool for early diagnosis and risk assessment of cardiovascular diseases.

REFRENCES

[1] World Health Organization, “Cardiovascular diseases (CVDs),” 2021.

[2] D. Dua and C. Graff, "UCI Machine Learning Repository," University of California, Irvine, 2019.

[3] K. Polat and S. Güneş, "A hybrid approach to heart disease classification," *Expert Systems with Applications*, vol. 38, no. 7, pp. 7904–7910, 2011.

[4] V. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 1998.

[5] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[6] D. Saha *et al.*, "Heart disease prediction using logistic regression," in *Proc. 2025 Int. Conf. Computer, Electrical & Communication Engineering (ICCECE)*, Kolkata, India, 2025, pp. 1–6, doi: 10.1109/ICCECE61355.2025.10940073.

[7] B. Duraisamy, R. Sunku, K. Selvaraj, V. V. R. Pilla, and M. Sanikala, "Heart disease prediction using support vector machine," *Multidisciplinary Science Journal*, vol. 6, p. 2024ss0104, 2023, doi: 10.31893/multiscience.2024ss0104.

[8] U. M. S. and L. Joshitha, "Heart disease prediction using decision tree variants: A comparative analysis," in *Proc. 2025 Int. Conf. Computing and Communication Technologies (ICCCT)*, Chennai, India, 2025, pp. 1–6, doi: 10.1109/ICCCT63501.2025.11019035.

[9] R. R. Devi, P. Dharshini, R. Hemala, and D. Swetha, "Heart disease prediction using random forest classifier," in *Proc. 2023 Int. Conf. Data Science, Agents & Artificial Intelligence (ICDSAAI)*, Chennai, India, 2023, pp. 1–8, doi: 10.1109/ICDSAAI59313.2023.10452459.

[10] S. Ghouhar Taj and K. Kalaivani, "Heart disease prediction using KNN algorithm approach," *ARPN Journal of Engineering and Applied Sciences*, vol. 19, no. 12, pp. 754–764, Jun. 2024.

[11] A. A. Shinde, S. N. Kale, R. M. Samant, A. S. Naik, and S. A. Ghorpade, "Heart disease prediction system using multilayered feed-forward neural network and back-propagation neural network," *International Journal of Computer Applications*, vol. 166, no. 7, pp. 32–36, May 2017, doi: 10.5120/ijca2017914080.

[12] N. Sabri *et al.*, "HeartInspect: Heart disease prediction of an individual using naïve Bayes algorithm," in *Proc. IEEE 11th Conf. Systems, Process & Control (ICSPC)*, Malacca, Malaysia, 2023, pp. 350–354, doi: 10.1109/ICSPC59664.2023.10420149.

[13] F. D. M. Inamdar, N. JayaKumar, and B. Bari, "Heart disease predictive analysis using association rule mining," in *Intelligent Communication Technologies and Virtual Mobile Networks*, G. Rajakumar *et al.*, Eds. Singapore: Springer, 2023, pp. 1–10, doi: 10.1007/978-981-19-1844-5_10.

[14] A. El-Ibrahimi, O. Daanouni, Z. Alouani, O. El Gannour, S. Saleh, B. Cherradi, and O. Bouattane, "Fuzzy-based system for coronary artery disease prediction using subtractive clustering and risk factor data," *Intelligence-Based Medicine*, vol. 11, 2025, doi: 10.1016/j.ibmed.2025.100208.

[15] M. Bhuiyan, C. Wah, N. Kamarudin, N. Ismail, and A. F. Ab. Nasir, "Heart disease prediction using case-based reasoning (CBR)," *Journal of Theoretical and Applied Information Technology*, vol. 102, no. 20, pp. 7275–7286, 2025.

[16] A. Diwan, S. Karmakar, and S. Soni, "Design and development of a machine learning-based hybrid model using BPNN and associative classifier for heart disease prediction," *MSW Management Journal*, vol. 35, no. 2, 2025, doi: 10.7492/vkf74q14.

[17] J. Kaur and B. Khehra, "Fuzzy logic and hybrid-based approaches for the risk of heart disease detection: State-of-the-art review," *Journal of the Institution of Engineers (India): Series B*, vol. 103, no. 2, pp. 681–697, 2021.

[18] M. Aghamohammadi, M. Madan, J. K. Hong, and I. Watson, "Predicting heart attack through explainable artificial intelligence," in *Proc. Computational Science – ICCS*, 2019, vol. 11537, pp. 633–645.

[19] N. Ziasabounchi and I. Askerzade, "ANFIS-based classification model for heart disease prediction," *International Journal of Electrical & Computer Sciences*, vol. 14, 2014.

[20] M. A. M. Abushariah, A. A. M. Alqudah, O. Y. Adwan, and R. M. M. Yousef, "Automatic heart disease diagnosis system based on artificial neural network and adaptive neuro-fuzzy inference systems," *Journal of Software Engineering and Applications*, vol. 7, pp. 1055–1064, 2014.

[21] A. S. Kumar, "Diagnosis of heart disease using fuzzy resolution mechanism," *Journal of Artificial Intelligence*, vol. 5, no. 1, pp. 47–55, 2011.

[22] A. Diwan "Design and Development of a Machine Learning based Hybrid Model using BPNN and Associative Classifier for Heart Disease Prediction", MSW MANAGEMENT -Multidisciplinary, Scientific Work and Management Journal ISSN: 1053-7899 Vol. 35 Issue 2, 2025, Pages: 1655-1665