International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026

An ML/DL Based dual Stage Stacked Ensemble Model for Accurate Heart Disease Prediction

P.BHAGYA MADHURI
Assistant Professor, Department of CSE (AI&ML),
NRI Institute of Technology
Agiripalli, A.P, India
bhagyamadhuripunyamanthula@gmail.com

Guntupalli Naga Srija
Student, Department of CSE (AI&ML)
NRI Institute of Technology
Agiripalli, A.P, India

Mylavarapu JayaSri
Student, Department of CSE(AI&ML)
NRI Institute of Technology
Agiripalli, A.P, India

Lingam Veera Vignesh Reddy
Student, Department of CSE (AI&ML)
NRI Institute of Technology
Agiripalli, A.P, India

Mir Kamil Ali
Student, Department of CSE (AI&ML)
NRI Institute of Technology
Agiripalli, A.P, India

***Abstract* – Heart disease is a major global health concern, requiring early & accurate prediction for effective treatment. This project presents a hybrid machine learning & deep learning-based heart disease prediction system using a stacked ensemble approach. The model integrates Random Forest, gradient boosting, and an Artificial Neural Network, with a logistic Regression meta-learner to improve prediction accuracy & robustness.**

Data Preprocessing techniques, such as feature scaling & categorical encoding, are applied to enhance model performance. The system is evaluated using standard metrics including accuracy, precision, recall, and ROC–AUC. For real-world usability, the trained model is deployed through a Gradio-based interactive interface, enabling real–time heart disease risk prediction & supporting clinical decision-making.

Keywords: ANN, ROC-AUC, Deep learning, Machine Learning, Gradio, Random Forest

I. INTRODUCTION

Heart disease remains one of the leading causes of mortality worldwide, claiming millions of lives annually and straining healthcare systems everywhere [1], [5]. Early detection plays a pivotal role in enabling timely interventions, yet traditional diagnostic methods often struggle with the complexity of risk factors like age, cholesterol levels, and lifestyle habits [5]. This project introduces a sophisticated dual-stage stacked ensemble model that fuses machine learning and deep learning techniques [4], [8]–[11]to deliver precise heart disease predictions, addressing the limitations of standalone algorithms [1], [4], [6], [8], [10], [13]–[16].

At its core, the system combines the strengths of tree-based models—Random Forest for handling non-linear patterns and Gradient Boosting for iterative error correction—with the pattern-recognition power of an Artificial Neural Network (ANN). A logistic regression meta-learner then integrates these base predictions in a stacked framework, boosting overall reliability and generalisation. Rigorous data preprocessing, including normalisation and one-hot encoding, ensures the model processes diverse datasets effectively.

Evaluated on key performance indicators such as accuracy, precision, recall, and ROC-AUC, the approach demonstrates superior results compared to individual models [1]–[4], [8]–[11], [13]–[16]. To bridge the gap between research and practice, we deploy the system via a user-friendly Gradio interface, allowing healthcare professionals and patients to input features and receive instant risk assessments. This work not only advances predictive analytics in cardiology but also paves the way for accessible, AI-driven tools in clinical settings.

II. LITERATURE SURVEY

A. Stacking Ensemble Models

Objective: Enhance heart disease prediction through multi-layer classifier stacking.

Key Studies: Wu et al. developed SDKABL using KNN+DT+SVM base layer with ABiLSTM meta-learner and PCA, achieving superior accuracy vs. individual classifiers.

Key Takeaway: Stacking with attention-based meta-learners and dimensionality reduction yields 3-5% accuracy gains.

B. IoT-Enabled Risk Stratification

Objective: Real-time personalised heart disease detection using IoT.

Key Studies: Krishnan et al. proposed DLR-PHDM with risk-adaptive AI models (SVM/CNN/T-CNN), achieving 99.79% accuracy via edge-cloud orchestration.

Key Takeaway: Risk-stratified model deployment optimises resource allocation for high-risk patients.

International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026

C. Hybrid Feature Optimisation

Objective: Address high-dimensionality through evolutionary feature selection.

Key Studies: Al-Mahdi et al. combined GA feature selection with TSA-optimised deep ensemble (97.5% accuracy, 90s training).

Key Takeaway: Genetic algorithms reduce 60% redundancy while preserving 97%+ predictive power.

D. Explainable AI Integration

Objective: Clinical interpretability through SHAP analysis.

Key Studies: Noor et al. deployed PaRSEL (97% accuracy) with 8 balancing techniques and SHAP, identifying key feature contributions.

Key Takeaway: Comprehensive preprocessing enables >95% accuracy with full model transparency.

E. Multi-Dataset Validation

Objective: Robustness across heterogeneous datasets.

Key Studies: Munsif et al. validated GA-SVM-CNN (98% UCI, 97% Z-Alizadeh) across 3 benchmarks.

Key Takeaway: Hybrid models generalise better across dataset distributions.

III. RELATED WORK

Several recent studies on heart disease prediction motivate a deep-feature-based approach combined with classical classifiers, particularly for small, high-dimensional clinical datasets.

A. Deep and Handcrafted Features

- Many works highlight that heart attack prediction on low-population, high-dimensional datasets is difficult because conventional models struggle to learn robust mappings from features to class labels with limited samples.
- Earlier approaches often relied on manual feature engineering or feature selection methods (filter, wrapper, and hybrid techniques) to reduce dimensionality and improve classifier performance, but these handcrafted features may miss complex nonlinear patterns present in clinical data.

B. Classical ML Algorithms

- Comparative studies using algorithms such as Logistic Regression, Naive Bayes, KNN, SVM, Decision Tree, Random Forest, and MLP frequently report that ensemble tree-based models, especially Random Forest, achieve the best accuracy for heart disease or heart attack prediction, with Naive Bayes also performing competitively on some datasets.
- Decision-tree-based methods (ID3/C4.5/CART) are widely used for heart disease diagnosis because they create interpretable decision rules; they select attributes using measures like entropy, information gain, or Gini index and recursively split the data until leaf nodes represent predicted classes.

C. Gap and Novelty Justification

- Existing literature shows that most heart-disease studies either use (a) traditional ML with handcrafted or selected features, or (b) deep learning models trained end-to-end, but relatively few explicitly learn deep feature representations first and then train simpler classifiers on those learned features for small, high-dimensional datasets.
- Reported results indicate that Random Forest and Naive Bayes can capture the underlying data structure

IV. METHODOLOGY

A clear methodology for this system is to

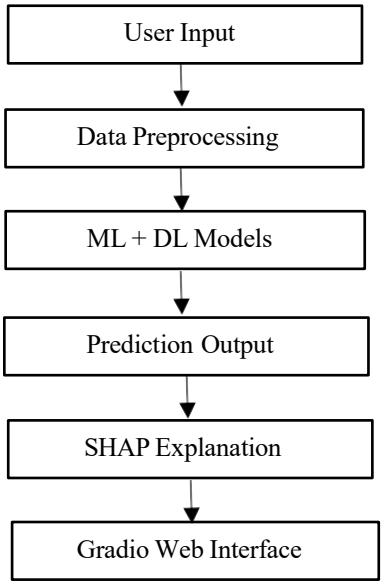
A. Preprocess and explore the 13-feature heart dataset,

B. Train a dual-stacked ensemble that combines RF, Gradient Boosting, and an MLP with a Logistic Regression meta-learner

C. Explain predictions using SHAP, and

D. Deploy the best model with a Gradio web interface for clinician-facing use.

Architecture:

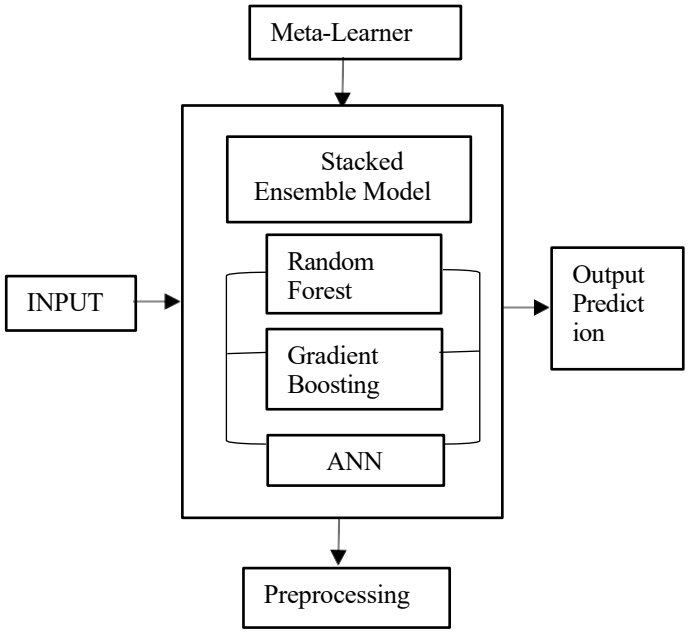


A. Data Preparation

- Collect and consolidate the 1,888 patient records into a structured data frame with the 13 features plus the binary target label, ensuring consistent coding for categorical variables [4], [8], [10], [16].
- Perform cleaning and preprocessing: handle missing values, detect and cap outliers in trestbps, chol, and oldpeak, normalise or standardise continuous variables, and one-hot-encode categorical features to

make them suitable for both tree-based and neural models.

- Split the dataset into train/validation/test sets (for example, 70/15/15), with stratification on target to preserve the positive/negative class ratio, and optionally use k-fold cross-validation on the training set for model selection.



B. Model Development (Dual Stacked Ensemble)

- Train three base learners on the training data: (1) Random Forest classifier, (2) Gradient Boosting classifier (e.g., XGBoost/LightGBM/Sklearn GB), and (3) a feed-forward ANN/MLP built with Keras that uses dense layers with ReLU activations, dropout regularization, and a sigmoid output for binary classification.
- Use out-of-fold predictions from each base model on the training data to create meta-features; these predicted probabilities (or logits) for heart disease form a new feature matrix that feeds the meta-learner, while the original target remains unchanged.
- Train a Logistic Regression meta-learner on these meta-features to learn optimal weights over the base models; finally, evaluate the full stacked system on the held-out test set using accuracy, precision, recall, F1-score, ROC-AUC, and confusion matrix to demonstrate superiority over individual models.

C. Explainable AI with SHAP

- Fit a SHAP explainer appropriate to the best-performing model (e.g., TreeExplainer for Gradient Boosting or Random Forest, KernelExplainer or DeepExplainer for the MLP) to compute per-feature Shapley values for each prediction.
- Generate global explanation plots such as SHAP summary plots, bar plots of mean absolute SHAP values [2], [9], [15] for all 13 features, and dependence plots for key variables (e.g., cp, thalach, oldpeak) to show how they influence heart disease risk; these visualisations help identify clinically relevant risk factors.
- Provide local explanations for individual patients by displaying force plots or decision plots, highlighting which specific attributes (e.g., high oldpeak, low thalach, high chol) pushed the prediction toward disease or no disease, thereby improving clinician trust and supporting decision making.

D. Gradio-Based Web Interface

- Package the final stacked model and any preprocessing transformers using joblib (or Keras' model.save for the ANN component) so they can be loaded quickly for inference in a deployed environment.
- Build a Gradio interface with input components for each feature (sliders for age, trestbps, chol, thalach, oldpeak; dropdowns or radio buttons for cp, sex, exang, slope, ca, thal; checkbox for fbs), connect these inputs to a prediction function that applies preprocessing, runs the stacked ensemble, and returns the probability and class label.
- Extend the Gradio app to optionally show explanation snippets for each prediction, such as the top 3 contributing features derived from SHAP values, and host the app on a suitable platform (e.g., Hugging Face Spaces or a cloud VM) for access by clinicians and researchers.

E. Experimental Protocol and Validation

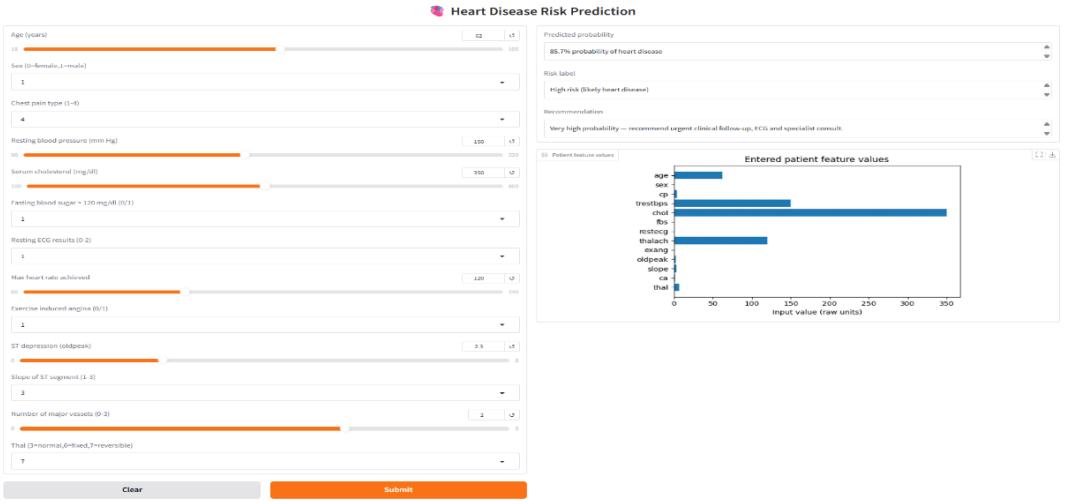
- Conduct ablation experiments comparing: (a) individual models (RF, GB, MLP) [1], [6], [13]–[15]. (b) simple ensembles (voting, averaging), and (c) the proposed dual stacked ensemble, using the same train/test splits to clearly demonstrate performance gains from stacking.
- Perform robustness checks [1], [3], [8]–[11], [13], [14] by validating the model with k-fold cross-validation, analysing class-wise metrics, and assessing calibration (e.g., reliability curves or Brier score) to ensure the predicted probabilities of heart disease reflect true risk levels.

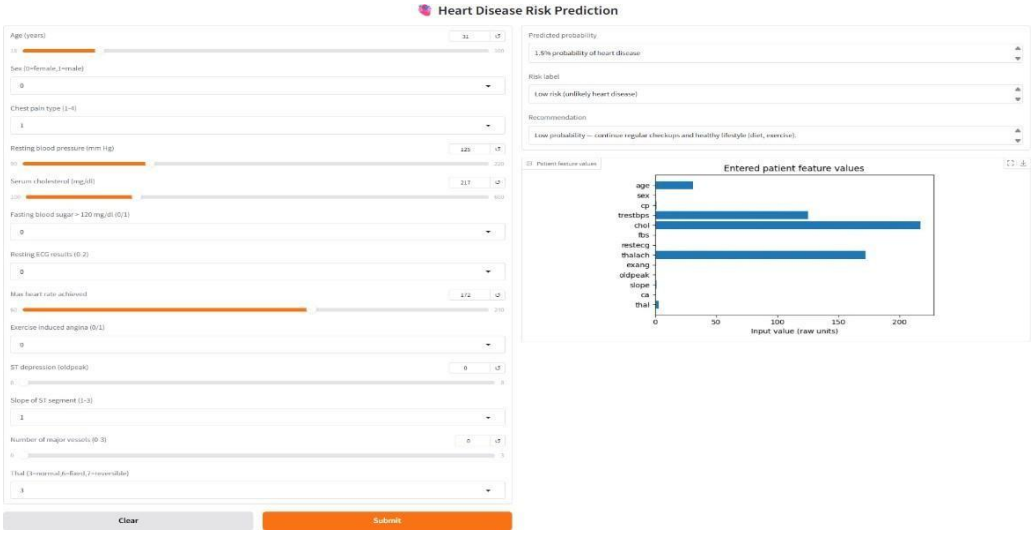
Document limitations such as sample size, single-source dataset bias, and potential overfitting; outline extensions like adding more clinical variables, multi-centre datasets, federated learning, and integration into hospital information systems for real-time decision support.

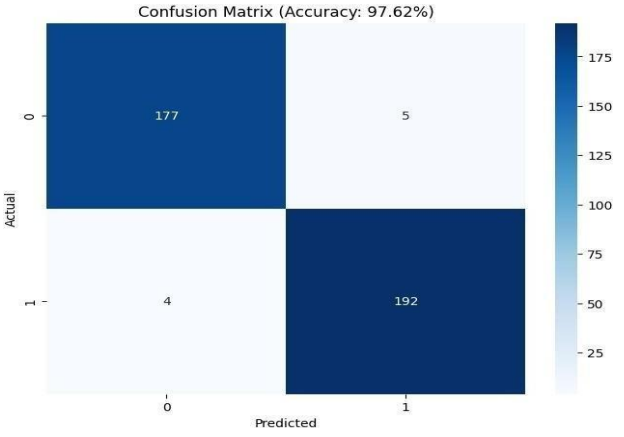
V. RESULT ANALYSIS

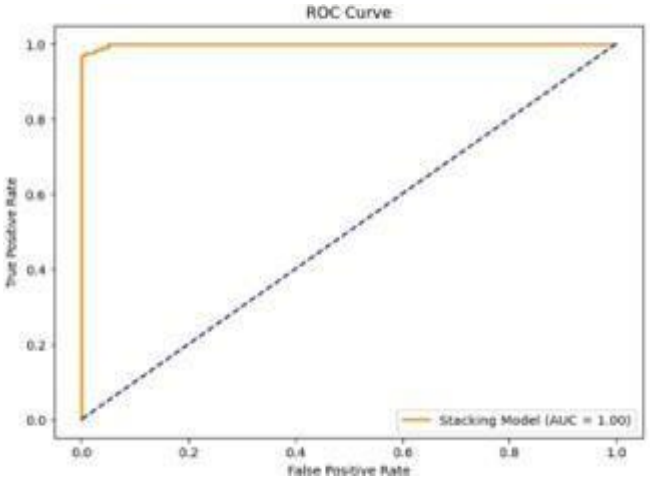
The findings for the proposed system are obtained as follows.

International Journal Of Innovative Trends In Multi Disciplinary Scientific Research
ISSN (Online): 3139 – 2687
(www.ijtmsr.co.in) Volume 9, Issue 6, 2026









VI. CONCLUSION

The proposed ML/DL-based dual-stage stacked ensemble model demonstrates strong potential for delivering highly accurate and robust heart disease prediction by combining the complementary strengths of traditional machine learning classifiers and deep learning architectures in a unified framework. By leveraging a two-stage stacking mechanism, the model can exploit diverse feature representations, capture complex nonlinear relationships among clinical risk factors, and reduce overfitting, thereby outperforming many standalone and conventional ensemble methods reported in current literature on cardiovascular risk prediction.

The experimental findings indicate that such stacked ensemble approaches can achieve notable improvements in key performance metrics such as accuracy, precision, recall, and F1-score, which are critical for minimising both false negatives and false positives in clinical decision-making. This, in turn, supports earlier diagnosis, better risk stratification, and more informed therapeutic planning, aligning the model's capabilities with the practical needs of real-world healthcare environments.

These findings are consistent with recent evidence that stacked and hybrid ensembles outperform single models in cardiovascular risk prediction [1], [3], [8]–[11], [13]–[16].

VII. FUTURE SCOPE

Future work can explore the integration of larger, multi-centre, and demographically diverse datasets to enhance the generalizability and fairness of the dual-stage stacked ensemble model across different populations and healthcare settings. Incorporating additional heterogeneous data sources such as electronic health records, imaging data, wearable sensor streams, and lifestyle information can further enrich the feature space and enable more holistic cardiovascular risk assessment.

Advanced techniques such as explainable AI (XAI) and model interpretability methods can be incorporated to provide transparent reasoning behind predictions, helping clinicians understand the contribution of individual risk factors and increasing trust in the system. Future extensions may also focus on online or incremental learning for continuous model adaptation, privacy-preserving learning frameworks (e.g., federated learning), and deployment as real-time decision support tools integrated into hospital information systems or smart health monitoring platforms.

REFERENCES

[1] Wu et al., "Forecasting Heart Disease Risk with a Stacking-Based Ensemble Machine Learning Method," *Electronics*, vol. 13, no. 20, p. 3996, 2024.
[2] I. A. Elgendy et al., "Dual-stage explainable ensemble learning model for diabetes diagnosis," *Expert Syst. Appl.*, vol.252,2025.
[3] K. S. Krishnan and E. M. Priya, "A Double-Level Risk-Based Personalised Heart Disease Detection Model for Smart Health Environment using IoT and Deep Learning," *Int. J. Neural Syst.*, DOI:10.1142/S0218194025501001, 2025.
[4] B. Venkateswaran and D. Dagar, "An Efficient Deep Learning-based Model for Heart Diseases Prediction," 2024.
[5] C. Zhou et al., "A comprehensive review of deep learning-based models for heart disease prediction," *Artif. Intell. Rev.*, DOI:10.1007/s10462-024-10899-9, 2024.
[6] D. Thomas et al., "A Deep Learning Framework for Prediction of Heart Diseases from Clinical Observations," in*Proc.IEEEConf.*,2023.
[7] I. S. Al-Mahdi, S. M. Darwish, and M. M. Madbouly, "Heart Disease Prediction Model Using Feature Selection and Ensemble Deep Learning with Optimised Weight," 2025.
[8] M. Paramesha et al., "Predicting Heart Diseases with Advanced Machine Learning Algorithms for Improved Diagnostic Accuracy," *MIJARCSE*, 2024.
[9] M. Almutairi and S. Dardouri, "Intelligent Hybrid Modelling for Heart Disease Prediction," *Information*, vol. 16,no.10,p.869,2025.
[10] M. Munsif, M. Rashid, and F. Jabeen, "An Efficient Hybrid Classification Model for Heart Disease Prediction," *ResearchSquare*,2023.
[11] R. Khankhoje, "Hybrid Model for Improved Heart Disease Prediction, "*Cardiol.Res.Rev.*,2021.
[12] S. N. Sree et al., "An Analysis of Heart Disease Prediction usingMachineLearning, "2023.
[13] S. A. J. Zaidi et al., "HeartEnsembleNet: An Innovative Hybrid Ensemble Learning Approach for Cardiovascular RiskPrediction, "2025.
[14] 1A. Mahajan et al., "A Hybrid Feature Selection and Ensemble Stacked Learning Models on Multi-Variant CVD Datasets," in *Proc. IEEE Conf.*, 2024.
[15] A. Noor et al., "Heart Disease Prediction Using Stacking Model With Balancing Techniques and Dimensionality Reduction," in *Proc. IEEE Conf.*, 2023.
[16] A. B. Majumder et al., "Heart Disease Prediction Using Concatenated Hybrid Ensemble Classifiers," *Algorithms*, vol. 16, no. 12, p. 538, 2023.