Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

Heart Disease Risk Prediction Using TabPFN, MLP and PSO Based Ensemble

Gannamani Keerthi Chowdary, Mounish Reddy
School of Computer Science and Engineering, VIT-AP University, Amaravati, India

***Abstract* - Cardiovascular disease is a leading cause of global mortality, underscoring the need for reliable early risk detection systems. This paper proposes a multi-class heart disease risk prediction framework that classifies patients into four risk levels: no risk, low risk, medium risk, and high risk. The system integrates TabPFN, a pre-trained transformer-based model for tabular data, with a Multi-Layer Perceptron (MLP) in a weighted ensemble. Particle Swarm Optimization (PSO) is employed to determine the optimal weight distribution between the two models. Evaluated on a dataset of 1,025 patient records from the UCI repository, the proposed ensemble achieves 98.54% classification accuracy, an AUC of 0.9989, precision of 0.9927, recall of 0.9840, and an F1-score of 0.9881. Results demonstrate that the PSO-optimized ensemble consistently outperforms individual models, confirming the value of combining complementary deep learning architectures for clinical risk stratification.**

Index Terms-Heart Disease Prediction, TabPFN, Multi-Layer Perceptron, Particle Swarm Optimization, Ensemble Learning, Multi-class Classification, Risk Stratification.

I. INTRODUCTION

Cardiovascular disease remains one of the leading causes of mortality worldwide, necessitating the development of reliable and automated tools for early risk detection. Conventional diagnostic methods depend heavily on clinical expertise and manual interpretation of patient data, processes that are inherently time-intensive and susceptible to human error. Machine learning has emerged as a compelling alternative, offering the ability to analyze large volumes of medical data and identify complex patterns that may elude traditional approaches.

Most existing machine learning solutions for heart disease prediction frame the problem as a binary classification task-simply determining whether or not a patient has heart disease. While useful, this binary output provides limited actionable information for clinicians who must assess not just the presence but also the severity of the condition. A more granular risk stratification would allow healthcare providers to better prioritize intervention strategies and personalize patient care.

This paper presents a multi-class heart disease risk prediction system that classifies patients into four distinct risk levels: no risk, low risk, medium risk, and high risk. The proposed system integrates two advanced machine learning models-TabPFN, a transformer-based model pre-trained on synthetic tabular datasets, and a Multi-Layer Perceptron (MLP)-within a weighted ensemble framework. The ensemble weights are optimized using Particle Swarm Optimization (PSO), a population-based metaheuristic algorithm, ensuring that each model's contribution is calibrated for maximum predictive accuracy.

The remainder of this paper is organized as follows. Section II reviews related work. Section III describes the dataset and methodology. Section IV presents experimental results. Section V concludes the paper and outlines future directions.

II. RELATED WORK

Extensive research has been conducted on applying machine learning to heart disease prediction. Kavitha and Suresh [3] applied traditional classifiers - Logistic Regression, Support Vector Machines (SVM), and Random Forest-to publicly available heart disease datasets, reporting moderate accuracy levels. While these models are interpretable and computationally efficient, they are largely limited to binary classification and do not generalize well to complex, nonlinear feature interactions.

Khan et al. [4] explored deep learning architectures, including Multi-Layer Perceptrons and Convolutional Neural Networks, for cardiac event prediction. Their results demonstrated improved accuracy compared to traditional methods; however, the black-box nature of deep learning models and their need for large labeled datasets remain persistent challenges in clinical settings.

Hollmann et al. [1] introduced TabPFN, a transformer-based model pre-trained on synthetic tabular data that treats classification as a probabilistic inference problem. TabPFN achieves competitive or superior performance compared to gradient boosting methods on small-to-medium datasets while requiring minimal task-specific training. Ye et al. [5] and Liu and Ye [7] further examined TabPFN's capabilities and extended its applicability to larger and noisier datasets.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

In the domain of ensemble learning, combining diverse models consistently improves prediction robustness. Ruiz-Villafranca et al. [8] applied a weighted fusion ensemble incorporating TabPFN in industrial IoT intrusion detection, demonstrating the benefit of metaheuristic-based weight optimization. PSO has also been widely used for hyperparameter tuning and model weight optimization due to its efficiency in large search spaces [3][4].

Despite these advances, a gap remains in the literature regarding multi-class risk stratification for heart disease using an ensemble of deep learning and transformer-based models with metaheuristic optimization. The present work addresses this gap directly.

III. METHODOLOGY

A. Dataset

The dataset used in this study is sourced from the UCI Machine Learning Repository [2] and comprises 1,025 patient records. Each record contains 13 clinical features: age, sex, chest pain type (cp), resting blood pressure (trestbps), serum cholesterol (chol), fasting blood sugar (fbs), resting ECG results (restecg), maximum heart rate (thalach), exercise-induced angina (exang), ST depression (oldpeak), slope of peak exercise ST segment, number of major vessels (ca), and thalassemia type (thal). The original binary target variable indicates the presence of heart disease.

B. Risk Level Generation

To transform the binary classification problem into a four-class risk stratification task, a clinically-informed target variable, risk_level, was derived from three medically significant features: oldpeak, trestbps, and chol. Patients without heart disease were assigned risk level 0 (no risk). Among those with heart disease (target = 1), low risk (level 1) was assigned when oldpeak < 1.0, trestbps < 140, and chol < 250; medium risk (level 2) when moderate elevation was present in any of these features; and high risk (level 3) when any feature exceeded critical thresholds (oldpeak ≥ 2.5, trestbps ≥ 160, or chol ≥ 300). The resulting class distribution was: no risk (499), medium risk (234), low risk (192), and high risk (100).

C. Preprocessing

Features were standardized using StandardScaler, transforming all values to zero mean and unit variance. Label encoding was applied to the risk_level target. The dataset was split into training (60%), validation (20%), and test (20%) subsets using stratified sampling to preserve class distributions across all splits.

D. TabPFN Model

TabPFN (Tabular Prior-Data Fitted Network) [1] is a transformer-based model pre-trained on a large number of synthetic tabular datasets. It models classification as a probabilistic inference problem, using attention mechanisms to capture global feature interactions without requiring gradient-based training from scratch. In this work, TabPFN was initialized with n_estimators = 2, float32 inference precision, and a fixed random seed. Fine-tuning was performed for 5 epochs using SGD (lr = 5×10^{-5}) through a custom meta-dataset dataloader.

E. MLP Model

A Multi-Layer Perceptron was implemented in PyTorch with two hidden layers of 128 and 64 neurons, ReLU activations, and dropout (rate = 0.2) for regularization. The model was trained for 50 epochs using the Adam optimizer (lr = 10^{-3}) with Cross-Entropy Loss. Softmax outputs were used to derive class probabilities for ensemble combination.

F. PSO-Weighted Ensemble

The ensemble combines the class probability outputs of TabPFN and MLP through a weighted average:

$$P\_ensemble = w_1 \times P\_TabPFN + (1 - w_1) \times P\_MLP$$

where w_1 is the TabPFN weight optimized by PSO. The PSO algorithm employed 30 particles over 50 iterations, with inertia weight w = 0.7 (decaying at 0.995 per iteration), and cognitive and social parameters $c_1 = c_2 = 1.4$. The fitness function maximized validation accuracy. The final predicted class was determined by argmax over the ensemble output probabilities.

IV. RESULTS AND DISCUSSION

A. Individual Model Performance

Table I summarizes the performance metrics of all models on the held-out test set (n = 205).

Model	Accuracy	AUC	F1-Score
TabPFN	98.05%	1.0000	0.9795
MLP	71.22%	0.9170	0.6917
Ensemble	98.54%	0.9989	0.9881

TABLE I. Performance Comparison of Models

TabPFN achieved 98.05% accuracy and a perfect AUC of 1.0 on the test set, confirming its superior capacity for tabular data classification through pre-trained contextual inference. The MLP reached 71.22% accuracy and AUC of 0.917, indicating meaningful but limited standalone performance on this four-class task. The ensemble surpassed both individual models with 98.54% accuracy, AUC of 0.9989, precision of 0.9927, and recall of 0.9840-validating the benefit of combining complementary models.

B. PSO Optimization

The PSO algorithm converged to optimal ensemble weights of $w_1 = 0.4967$ (TabPFN) and $w_2 = 0.5033$ (MLP), achieving a validation accuracy of 98.05%. The near-equal weight distribution indicates both models contribute meaningfully to the ensemble, with PSO effectively balancing their complementary strengths. Convergence was achieved within 50 iterations, demonstrating PSO's efficiency in this one-dimensional weight search space.

C. Risk-Level Classification Analysis

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 04 , April - 2026

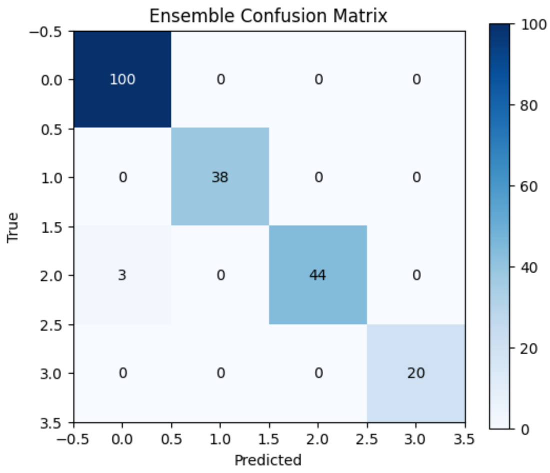

FIGURE 1: Ensemble Confusion Matrix

The ensemble confusion matrix reveals near-perfect classification across all risk levels. Risk Level 0 (no risk, n = 100): 100/100 correct. Risk Level 1 (low risk, n = 38): 38/38 correct. Risk Level 2 (medium risk, n = 47): 44/47 correct, with 3 misclassifications. Risk Level 3 (high risk, n = 20): 20/20 correct. The only misclassifications occurred between medium-risk and adjacent categories, which is expected given overlapping feature distributions.

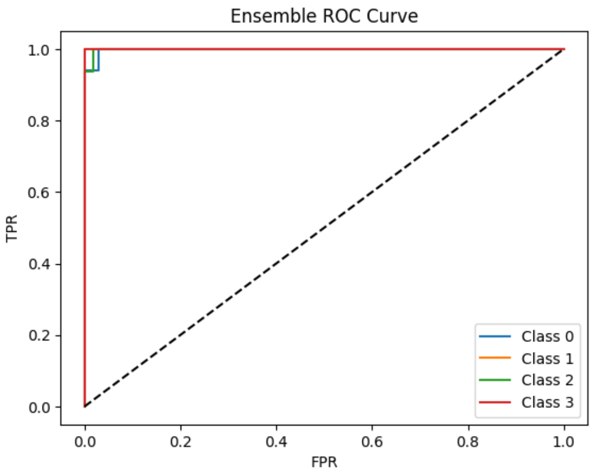

FIGURE 2: Ensemble ROC Curve

ROC curves for all four classes were clustered near the top-left corner, confirming strong discriminative ability across all risk categories.

D. Discussion

The results demonstrate that integrating TabPFN with MLP and calibrating ensemble weights via PSO yields highly accurate multi-class risk stratification for heart disease. The four-tier classification scheme provides richer clinical information than binary approaches, enabling more targeted intervention decisions. TabPFN's pretrained priors prove particularly effective for this structured medical dataset, while MLP contributes complementary nonlinear feature learning.

Key limitations include the relatively small dataset size (1,025 records), rule-based risk level thresholds that may not fully capture clinical complexity, and the limited interpretability inherent to both models. Future work should address these through clinician-validated risk thresholds, larger and more diverse datasets, and integration of explainability frameworks such as SHAP or LIME.

V. CONCLUSION

This paper presented a heart disease risk prediction system combining TabPFN and MLP within a PSO-optimized weighted ensemble. The system achieved 98.54% accuracy and an AUC of 0.9989 on a four-class risk stratification task, outperforming individual model baselines. The proposed multi-class framework advances beyond conventional binary approaches, offering granular risk assessment that can meaningfully support clinical decision-making. Future extensions will focus on larger datasets, clinical validation of risk thresholds, and deployment as a real-time clinical decision-support tool.

ACKNOWLEDGMENT

The authors thank Dr. Hussain Syed, Associate Senior Professor, VIT-AP University, for his invaluable guidance. We also acknowledge the UCI Machine Learning Repository for providing the heart disease dataset used in this study.

REFERENCES

[1] N. Hollmann, S. Müller, K. Eggensperger, F. Hutter, and M. Lindauer, "TabPFN: A Transformer That Solves Small Tabular Classification Problems in a Second," in Proc. ICLR, 2023.

[2] D. Dua and C. Graff, "UCI Machine Learning Repository," University of California, Irvine, 2019. [Online]. Available: https://archive.ics.uci.edu

[3] K. Kavitha and R. M. Suresh, "Prediction of Heart Disease Using Machine Learning Algorithms," Int. J. Sci. Technol. Res., vol. 9, no. 3, pp. 1023–1026, 2020.

[4] M. A. Khan et al., "A Hybrid Deep Learning-Based Model for Heart Disease Prediction," Comput. Biol. Med., vol. 136, 2021.

[5] H. J. Ye, S. Y. Liu, and W. L. Chao, "A Closer Look at TabPFN: Understanding Its Strengths and Extending Its Capabilities," arXiv preprint, 2022.

[6] L. Grinsztajn et al., "TabPFN-2.5: Advancing the State of the Art in Tabular Foundation Models," arXiv preprint, 2024.

[7] S. Y. Liu and H. J. Ye, "TabPFN Unleashed: A Scalable and Effective Solution to Tabular Classification Problems," arXiv preprint, 2023.

[8] S. Ruiz-Villafranca et al., "WFE-Tab: Overcoming Limitations of TabPFN in IIoT-MEC Environments with a Weighted Fusion Ensemble-TabPFN Model for Improved IDS Performance," 2023.