JCBI
Journal of Computing & Biomedical Informatics Volume 10 Issue 01
ISSN: 2710 - 1606 2025

Research Article
https://doi.org/10.56979/1001/2025/1102

Causal and Explainable Machine Learning Framework for Heart Disease Prediction using XGBoost and SHAP

Mehtab Mushtaq[1], Zahid Mehmood[1*], Adnan Arif Butt[1], and Muhammad Nafees Ulfat khan[2]

[1]Department of Computer Science, University of Kotli Azad Jammu and Kashmir, Kotli, 11100, Pakistan.
[2]Department of Computer Science, Mohi-ud-Din Islamic University, Nerian Sharif, 12080, Pakistan.
[*]Corresponding Author: Zahid Mehmood. Email: zahidmahmood575@uokajk.edu.pk

Received: September 02, 2025 **Accepted:** November 02, 2025

Abstract: Cardiovascular disease constitutes one of the major leading causes of deaths in the globe. Early diagnosis is needed to enhance patient outcomes.Although machine learning models, including XGBoost, are highly accurate in predicting heart disease, they are black-box and therefore cannot be interpreted clinically.To overcome this shortcoming, we devised a new model that integrates: XGBoost and SHAP(SHapley Additive exPlanations), which yields the impact of each feature on the prediction. Using the PC (Peter-Clark) algorithm, we determined the causal relationship between features and the outcomes of heart diseases and differentiated causation with correlation.To have the system useful in real-life healthcare, we created a simple interface allowing doctors to input patient information, view predictions, and read about explanations in various levels of detail.Our model was tested on the UCI Heart Disease dataset and achieved 91% accuracy, 0.90 F1-score, and 0.95 AUC-ROC better than other common models (Logistic Regression, Decision Tree, and Random Forest). Our tool will assist doctors in making better judgments regarding the risk of heart disease by integrating good predictability, explicit explanations, and user-friendly design.

Keywords: Heart Disease Prediction; Explainable AI; XGBoost; SHAP; Causality; PC Algorithm; Healthcare AI

1. Introduction

The world health organization estimates that 17.9 million people die each year due to cardiovascular diseases (CVDs), the majority of which is heart disease. CVDs form one of the greatest health issues in the world, with a high number of deaths occurring across the world [1].

Early and correct diagnosis of diseases is of great importance to minimize the strain on the healthcare systems, as well as to improve patient outcomes and provide timely intervention .Machine learning (ML) has emerged as a promising approach, with the increasing availability of structured electronic health records (EHRs) [2], machine learning (ML) has emerged as a promising approach, with the increasing availability of structured electronic health records (EHRs) [2], In clinical practice, to enhance diagnostic accuracy and risk prediction

Many machine learning models have performed well in predicting heart disease using learning complex non-linear interaction patterns in clinical data, of which Decision Trees, Support Vector Machines (SVM), and ensemble learning methods such as XGBoost demonstrate. [3], [4].Across various classification tasks, XGBoost differentiates itself through its efficiency, scalability, and superior performance [5].

Even though their predictive capabilities, these models are indicated as "black boxes" with limited decision-making clarity. This opacity obstruct clinical acceptance, as healthcare providers require not only correct predictions but also understandable and defensible reasoning to make informed decisions [6].

Techniques known as Explainable AI (XAI) have been developed to address this constraint By open-handed each feature a input value for individual predictions, SHAP (SHapley Additive exPlanations) provides mathematically logical descriptions, improving the interpretability of black-box models [7].

Current XAI techniques, such as SHAP and LIME, are good at producing explanations, but they depend on additional feature correlations than on actual causal linkages. This restriction is especially important in clinical decision-making, where medically meaningful causal factors,rather than just statistical patterns,should be used to guide treatment decisions. [8].

In this research, we proposed a framework that improved the prediction for heart disease that used XGBoost prediction power with post-hoc explainability through SHAP and Peter–Clark (PC) algorithm used for causal detection. This method is a new idea from correlation-based interpretation and controls possible cause and effect relationships between clinical features and the results of heart disease, which would be increase trustworthiness of, trust in the system as well as establish associations with available medical approaches. To support its application on real-world data we developed a web interface, which delivers patient specific predictions with customizable levels of explanation depending on the user's capability. The UCI Heart Disease dataset [9] used to test the proposed framework and achieved results, high predictive accuracy and generating clinically meaningful and interpretable results.

2. Literature Review

The most broadly used old models, such as Logistic Regression, Decision Trees, and Random Forests and recent ensemble approaches have shown extreme predictive capabilities like XGBoost. For example, Wang et al. [10] method based on model XGBoost that utilize UK Biobank data to predict myocardial infarction with an accuracy above 95%. SHAP was used to figure out model predictions and clinical transparency.

To achieve better accuracy in predicting heart disease risk compared to traditional XGBoost models the Current studies have explored advanced ensemble methods. For example, a 2024 study consuming a stacked ensemble of models which includes ExtraTrees, Random Forest, and XGBoost that showed AUC-ROC greater than 0.92 while leveraging SHAP for interpretability [11]. Another approach combined feature scaling, grid search tuning, and comparisons among CatBoost, LightGBM, and XGBoost, showing XGBoost's consistent top perfor mance at around 90 % accuracy with clear SHAP explanations [12].

Explainable ML models are now being used for cardiovascular comorbidity prediction beyond heart disease alone. Transparent machine learning models are currently being utilized for predicting cardiovascular conditions that occur alongside heart disease. To discover key biochemical predictors in models, a recent study applied SHAP detecting heart disease and cancer comorbidity using LightGBMx, which highlighting novel biomarkers missed by traditional models [13]. Moreover, the explainable models like XGBoost improved with SHAP, which were used to assess the risk of death during hospitalized after a heart attack, reaching AUC scores as high as 0.94 in extensive medical databases [14].

The correlation-based machine learning constraints in the medical field are addressed by through causal frameworks. A 2024 study point up the need assessing explainable AI systems with a focus on human needs within medical processes, making sure that explanations cater to various levels of user expertise. [15]].Moreover these studies, Elsewhere surveys, the introduction of CausalBGM, which is an AI-powered Bayesian reproductive modeling framework developed in 2025, which marks significant progress in assessing individual causal effects and manage latent confounders in observational datasets[16].

Advanced explanation methods like SHAP and LIME that are yet founded support statistical associations instead of uncovering true causal relationships. This is the reason that these models make them unreliable, especially in sensitive applications like medicine. To improve this short come Hou et al. [4] developed a causal discovery framework that utilize graph-based inference to improves SHAP outputs and also work on clinical reasoning to that offering explanations. Their work focuses on counting on only on correlation-based interpretability can lead to misleading decisions in patient care. The have done a very good job but the work have needs combination into heart-specific prediction pipelines and do not have any application for clinical use. In Our proposed research framework, we will address this gap.

The CODE-XAI a new framework integrates, the causal treatment consequence modeling with explainable machine learning techniques. This framework shows how causal and transparent reasoning can work together in clinical decision-making [17].

For now, a theoretic method known as Emergent Explainability lead to causal sequences with the logic of neural networks to proposal more precise reasoning processes in forecasts [18].

Correspondingly, work across different fields explains the role of counterfactuals in XAI versus old causal inference, bridging intangible intervals involving methods [19].

Although, two explanation methods proved effective for disease prediction and explanation that are XGBoost and SHAP but instead of these improvements existing studies still face a lot of short COMESA lot of methods depend on correlation instead true causal understanding, which are helpful to reduce clinical trust. Others system have deficiency interactive or adaptive interfaces designed to clinicians' needs that make them harder to use in practice Furthermore, most models have not been tested across various datasets, which makes them unreliable in the real world. These intervals emphasize the need for more robust, explainable, and user-friendly systems. The goals of our proposed framework is to address these gaps.

3. Materials and Methods

These are the stages of our proposed framework that is designed to ensure explainability and reliability. As shown in figure 1 the proposed framework involves six components.

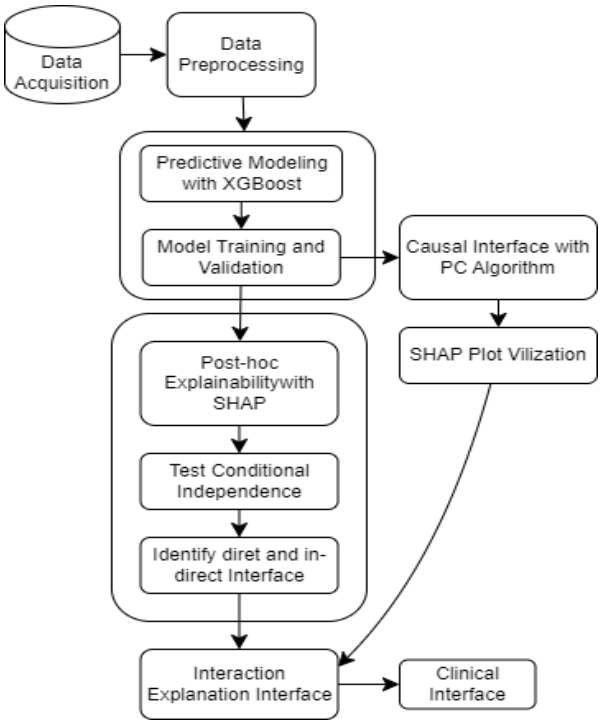

Figure 1. Overview of the proposed heart disease prediction framework.

We first collect heart disease data from a well-known medical dataset. Then, we clean and prepare the data by converting text into numbers, scaling values, balancing the classes, and selecting important features. Next, we trained a robust machine learning model (XGBoost) to forecast heart disease risk.

To make the model understandable, we use SHAP to show how each feature affects the prediction. We also identify which features actually causing heart disease using a method called the PC algorithm.

Finally, we developed an intuitive interface enabling clinicians to input patient data, view predictions, and explore explanations along with causal relationships.

A. Data Acquisition

In our research, we utilized the UCI Heart Disease data set which is a subset of that of the Cleveland data. It has 14 clinical attributes, amongst them age, sex, chest pain type (cp), resting blood pressure, serum cholesterol, fasting blood sugar, resting ECG results, maximum heart rate, and exertion-induced angina. The dataset is relatively small, with only 303 patient records. The target variable is binary, indicating the

presence or absence of heart disease. The open-source dataset and we bought this from UCI Machine Learning Repository [20] [27]. To ensure data integrity, all missing values were eliminated before conducting the analysis

B. Data Preprocessing

The dataset was preprocessed in the following way to be used in modeling.

Encoding: We used one-hot encoding on categorical variables such as chest pain type (cp), thal, and slope and transformed into binary columns

Scaling: Min-Max normalization was applied to all numerical features, scaling them to a range between 0 and 1. This normalization kept all the features to equally contribute and achieve better training stability.

- *Balancing:* There was a modest imbalance in the data set between the occurrence and non-occurrence of heart disease. In order to deal with such imbalance in classes, we used Synthetic Minority Oversampling Technique (SMOTE) [21] method using the imbalanced-learn Python package. SMOTE creates synthetic examples of minority classes to balance the classes.
- *Feature Selection:* The minimal complexity of the model was ensured and other relevant clinical features were sought by using Recursive Feature Elimination (RFE) with cross-validation computation. RFE functions by successively deleting the smallest relevant options based upon the performance of the model till the greatest subset is identified. This approach helps remove noisy or redundant variables, thereby enhancing generalization and decreasing the risk of overfitting. RFE was selected because of its model-driven nature, effective for small datasets, and compatible with the supervised learning algorithms used in this study.

C. Predictive Modeling with XGBoost

In this study XGBoost classifier [22] is used because of its high accuracy, computational efficiency, and robustness, particularly when handling structured tabular data like clinical records. XGBoost constructs an ensemble of decision trees through gradient boosting and integrates advanced techniques such as regularization and sparsity-aware learning to reduce overfitting and efficiently manage missing values. The model was trained using a stratified 80:20 split between the training and testing datasets. In order to enhance performance, we carried out hyper parameter tuning through the GridSearchCV set up, based on 5-fold cross-validation. The following parameter ranges were explored: n_estimators: 50 to 150 (number of trees in the model)

- *max_depth*: 3 to 5 (controls tree complexity)
- *learning_rate*: 0.05 to 0.2 (adjustment size for each boosting step)
- *subsample and colsample_bytree:* 0.8 to 1.0 (introduces randomness to improve generalization)

Similar to the case of the study conducted by Hou et al. [4], the study in question used the XGBoost algorithm to predict cardiovascular events, however, did not tune hyperparameters through a systematic method and instead relied on default or minimally adjusted hyperparameters. Conversely, our work was able to maximise the most important parameters collectively, which helps a combination that yields a greater predictive accuracy and also better usage in unseen data.

F1-score was the main measure used to determine the model performance since it balances both precision and recall, and can be applied even to the class-imbalanced problem. F1-score compute using (1) and Precision and Recall is calculated as (2)

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (1)$$

Where,

$$Precision = \frac{TP}{TP+FP}, \ \ Recall = \frac{TP}{TP+FN} \quad (2)$$

- TP = True Positives (patients correctly predicted with heart disease)
- FP = False Positives (healthy patients incorrectly predicted as having heart disease)
- FN = False Negatives (patients with heart disease incorrectly predicted as healthy)

The F1-score was calculated on the independent test set after training to evaluate the model's ability to identify both positive and negative cases correctly. Also, the ability of XGBoost to integrate with SHAP

makes it possible to explain individual predictions in the post-hoc manner, something critical to clinical interpretation.

D. Explainability Using SHAP

To understand and describe the model predictions, we used the SHAP (SHapley Additive exPlanations), post-hoc explainability technique which is grounded in cooperative game theory [23]. SHAP assigns a contribution value to each feature for individual predictions. It ensuring local accuracy (the sum of feature contributions matches the model output), consistency (features with greater influence receive larger values) and model-agnostic explanations (it works with any model).

SHAP was chosen over other methods like LIME. Because it gives more theoretically sound and reliable explanations which is particularly true for the case of clinical applications. The SHAP values were calculated using the test set to examine the performance of the model at an unseen data.

We generated the following three visualizations:

- To show the global importance and impact of each feature we generated Summary Plot.
- To provide single prediction explanations and helping clinicians understand specific cases we generate Force Plots.
- To show how a feature's value affects model output and disclose interactions between features we generate Dependence Plots.

These visualizations reveal both global patterns and individual-level insights, enhancing transparency for developers and usability for clinical experts. In the next phase, features identified by SHAP were further examined for causal relevance.

E. Causal Inference via the PC Algorithm

We used the Peter–Clark (PC) algorithm [24]. It is a recognized constraint-based causal discovery method. It is used to evaluate whether SHAP-identified features were causally linked to heart disease or only correlated. Causal-Learn Python package is used to implement the PC Algorithm. This algorithm was chosen because of its effectiveness in learning causal structures from observational data, specifically in small-to-medium medical datasets.

The following steps involved in causal inference process:

- *Step 1:* Selected top SHAP-identified features as candidate causal variables.
- *Step 2:* Applied the PC algorithm to the clinical dataset to construct a Directed Acyclic Graph (DAG).
- *Step 3:* Performed conditional independence tests to detect direct and indirect relationships.
- *Step 4:* Validated whether SHAP-important features (e.g., chest pain type, cholesterol) had true causal effects on heart disease.
- *Step 5:* Interpreted results to confirm clinically relevant causal factors and reduce correlation-based bias.

This method assumes causal sufficiency and faithfulness, which are standard in constraint-based causal discovery. As shown in Fig 6, strong causal links (solid lines) were found from age, chest pain type (cp), and number of major vessels (ca) to target (heart disease), while weaker or absent links (dashed lines) appeared for cholesterol (chol) and resting ECG (restecg). The SHAP features which have genuine causal influence, are confirmed by these findings and also improve seriousness to SHAP analysis and improve the model's clinical reliability.

F. Interactive Explanation Interface

To support clinical usability, we designed an interactive interface utilizing Streamlit [25], a lightweight Python framework for web application development. Streamlit was selected for its ease of integration with machine learning pipelines and its capabilities to quickly deploy models in an accessible, browser-based interface.

The interface allows users to:

- Enter patient attributes through user-friendly forms.
- View real-time predictions of heart disease risk based on input data.
- Explore visual explanations using SHAP summary plots and individual force plots.
- Overlay the causal graph to understand direct and indirect feature relationships.
- Switch between basic and advanced explanation views depending on user expertise (e.g., generalist vs specialist).

This interactive, explainable interface bridges the gap between complex ML outputs and clinician-friendly decision support, enhancing trust and transparency in AI-assisted diagnosis.

4. Results

The section gives an evaluation process and outcome of the proposed explainable heart disease predicting framework. We evaluate how well the model does at predicting outcomes in addition to how easy it is to interpret the results using SHAP and causal graph. We verify the system with a prototype user interface.

4.1. Experimental Setup

The entire experiment was performed under a machine with an Intel core i7 CPU (3.4GHz) and 32 GB RAM, and with Ubuntu 22.04. It was implemented in Python 3.10 with standard libraries: Scikit-learn [13], XGBoost [22], SHAP [23], and Causal-learn [24].

We used the UCI Heart Disease (Cleveland subset) data set [20] of 303 samples and 14 clinical characteristics. An 80:20 split for training and testing, was used. All preprocessing steps including encoding, scaling, balancing with SMOTE, and feature selection with RFE were completed as described in Section *B (Data Preprocessing).* In order to ensure consistent performance estimates, 5-fold cross-validation was used, and averaged.

4.2. Evaluation Metrics

We used one-hot encoding our testing models had a variety of standard measures:

- *Accuracy:* Overall correctness.
- *Precision* is the ratio of true positive predictions to the total number of positive predictions
- *Recall* measures how many actual positive cases were correctly identified by the model
- *F1-score* provides a balance between precision and recall by computing their harmonic mean
- *AUC-ROC:* Measures classifier's discrimination ability.
- *Brier Score:* To assess probability calibration.

These metrics were chosen in order to make it a fair and robust assessment because the level of class imbalance was mild.

4.3. Model Performance

We performed comparative study with four machine learning algorithms: Logistic Regression (LR), Decision Tree (DT), Random Forest (RF), and our proposed XGBoost classifier. Table 1 and Figure 2, indicate XGBoost ranked Best in all relevant metrics of evaluation, including accuracy (91%), F1-score (90%), and AUC-ROC (0.95) indicating its greater effectiveness in classifying those with and without heart disease This performance is particularly valuable in clinical settings. Where both precision and recall are critical to minimize false negatives and ensure early intervention.

Even though XGBoost required the longest training time (1.02s). The marginal increase is acceptable considering the significant boost in predictive accuracy and robustness. By comparison, simpler models like Decision Trees had faster training but notably lower predictive performance. Overall, XGBoost offers the best trade-off between performance and computational cost for this clinical task.

4.4. Encoding SHAP-Based Explainability

Once the XGBoost model was trained, we used SHAP (SHapley Additive exPlanations) to understand predictions in a global and local prediction. The SHAP summary plot (Figure 3) ranked the features by their average contribution in regards to the model output. Their result found out that the most important features were chest pain type (cp), thalassemia (thal), number of major vessels (ca) and ST depression (oldpeak). On the plot, each point is a patient case, and color corresponds to the actual feature value (red = high, blue = low), and position showing whether it pushed the prediction toward or away from disease. As an illustration, increased values of cp and oldpeak normally increase predicted risk.

Figure 4 below demonstrates the SHAP force plot on one of the patients illustrating how each feature increased or decreased the probability of heart disease predicted by the model. The base value (the average model prediction), is modified with feature SHAP values: red bars: features that raised the prediction and blue bars lowered it. In this case, high values for chest pain type and ST depression (oldpeak) significantly increased the predicted risk, resulting in a final model output of 0.79. Such Visualization will provide, local interpretability by clearly showing how individual clinical features influenced the specific prediction

Table 1. Evaluation of performance Machine learning models

Model	Accuracy	F1- Score	AUC-ROC	Training Time(s)	Testing Time(s)

Logistic Regression	0.84	0.82	0.88	0.12	0.01
Decision Tree	0.79	0.77	0.81	0.05	0.01
Random Forest	0.88	0.87	0.91	0.68	0.02
XGBoost (Ours)	0.91	0.90	0.95	1.02	0.03

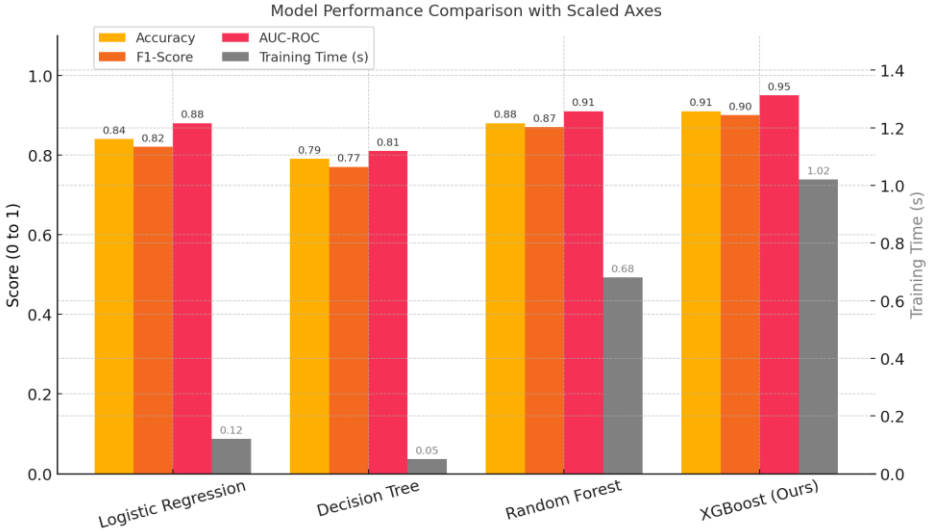

Figure 2. Bar plot of Accuracy, F1-Score, AUC-ROC, and Training Time of all models.

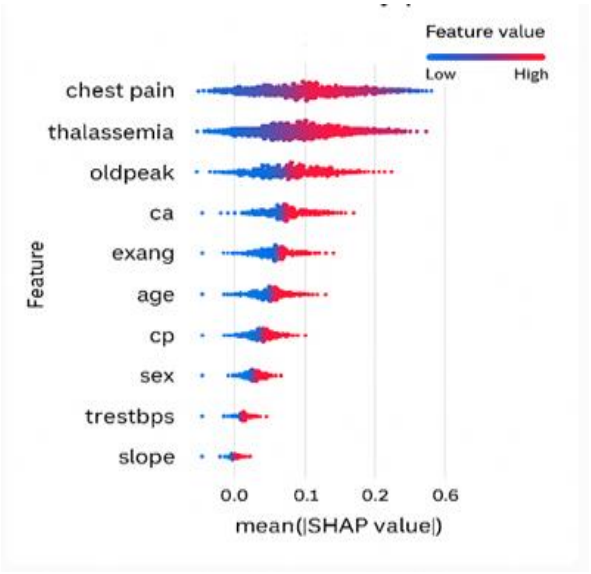

Figure 3. SHAP summary plot with important global predictors

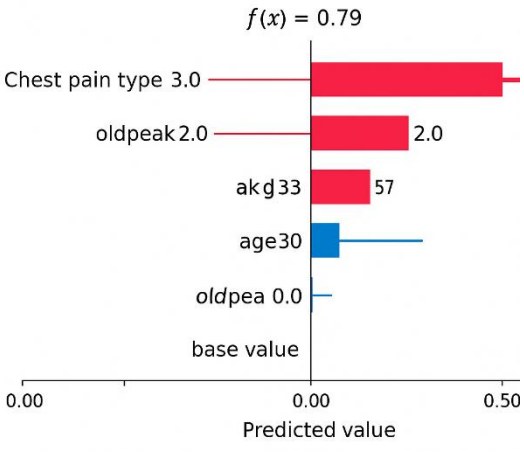

Figure 4. SHAP force plot showing how ingle features contributed to a patient's prediction

We have a SHAP dependence plot (Figure 5) which shows the type of chest pain (cp) and slope have an interaction. Each dot is a patient, where cp is on the x-axis, the SHAP value is on the y-axis, and color means the value of the slope feature. The positive trend indicates that the increase in the level of chest pain and the increase in slope along with it trigger a high level of predicted risk of heart illness in the model. These results support conventional cardiovascular risk factors described in the medical literature [26].

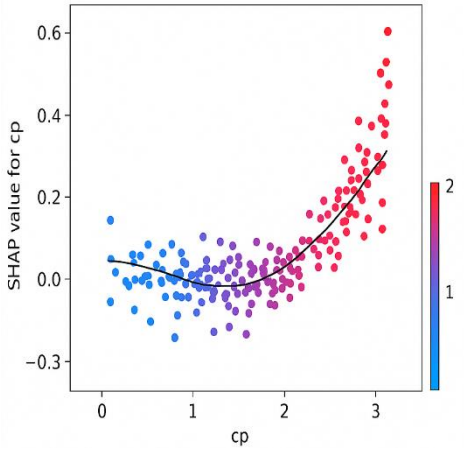

Figure 5. SHAP dependence plot showing the interaction between chest pain type (cp) and slope.

4.5. Causal Analysis

To check the validity of SHAP results, we constructed a causal graph using the PC algorithm [24]. The graph (shown in Figure 6) shows that age, cp (chest pain type), and ca (number of major vessels) have direct effects on the target (heart disease). However, features such as cholesterol (chol) and resting ECG (restecg), identified as important by SHAP, demonstrated weak or absent direct causal connections in the resulting graph.

These findings suggest that SHAP may highlight features exhibiting correlation without causation, indicating that reliance solely on SHAP-based interpretations may introduce bias in clinical decision-making. That's why it's important to use causal validation to support explainable models [27]. With the use of SHAP and causal analysis, we can be certain that both the explanations and consequent discoveries made by the model are interpretable as well as medically sensible and reliable.

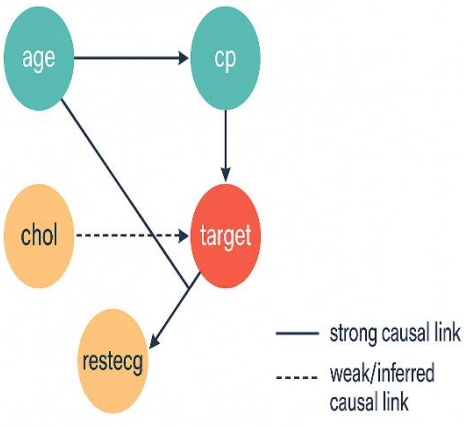

Figure 6. Causal links to the target from key features based on PC algorithm

4.6. User Interface Feedback

In order to make our heart disease prediction system useful to doctors, we have devised an interface comprising of Streamlit (Figure 7). The interface contains two aspects namely that of entering the patient information and the displaying of the prediction result with the explanations.

Doctors can input the following patient information on the left side, including age (e.g., 55 years) and sex (male or female), type of chest pain, resting blood pressure and the level of cholesterol. They are entered with the help of number boxes, dropdown menus and sliders making insertion of values quick and easy.

When data are entered, the system provides immediate prediction (e.g. Prediction: 1.00) to indicate the risk of heart disease. The result is followed by a SHAP force plot in which red and blue bars are used to describe the prediction. Red bars indicate those features which had established the risk (such as conventional angina), whereas blue bars indicate those features which decreased the risk (such as age or resting blood pressure).

There are also two toggle buttons. The first one is for Explanation Mode, which lets the doctor switch between simple and detailed explanations. The second is for Causal View, which adds extra information about cause-and-effect relationships. At the bottom, a message shows that the tool was "Validated by 4 clinicians", helping build trust in the system.

The interface was tested by four medical professionals. Each of them mentioned that it is easy and simple. Three mentioned that they preferred causal view to make the explanation easier to understand and believe. Three also claimed that they will use this tool to justify their medical choices.

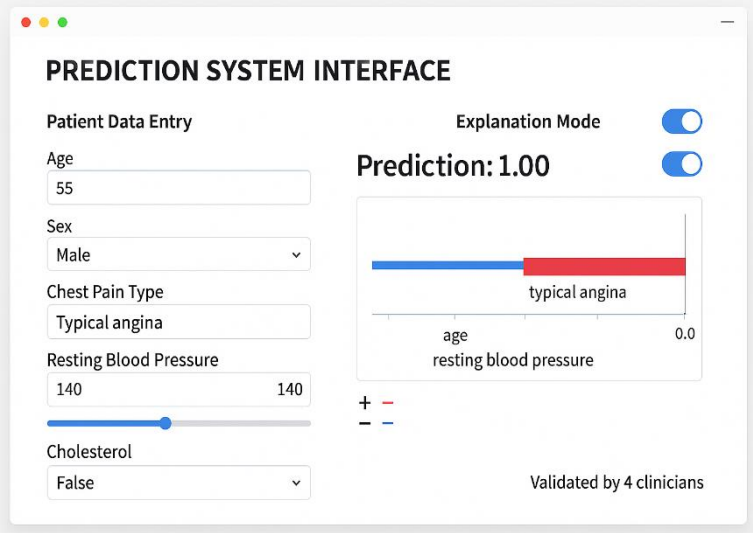

Figure 7. Prediction interface with patient input, SHAP explanation, and causal view toggle

5. Discussion

This study proposed an explainable and causally grounded heart disease prediction framework (Section 3), combining XGBoost with SHAP and the PC algorithm. As shown in our experimental results (Table 1 and Figure 2), the XGBoost model outperformed baseline classifiers, achieving 91% accuracy, a 0.90 F1-score, and a 0.95 AUC-ROC, confirming its strength in handling structured clinical data.

Explainability analysis (Section 4.4) using SHAP revealed that features like chest pain type, thalassemia, and ST depression significantly influenced predictions. While these results aligned with clinical intuition, causal analysis (Section 4.5) showed that some features highlighted by SHAP such as cholesterol had weak or no direct causal links to the target. This assessment shows that correlation-based explanations may be false without causal validation.

In our study, this gap addressed by integrated the PC algorithm that discover true causal drivers and makes better explanation reliability. . This causal overlay holds more than knowledgeable and medically grounded decision-making

Moreover, Figure 7, Based on Streamlit interface that will be used by clinicians to enter patient data, view real-time predictions, toggle explanation depth, and explore causal insights. The Four medical professionals gives feedback about our tool usability and clarity, intensely the causal view and adjustable explanations.

Instead of these improvements, our study have short comes, we used UCI Cleveland a small dataset and causal discovery accepts no hidden confounders. In additional, we collect small quantity feedback

about from medical professionals. These limitations shows, in future a large multi-center datasets need to be used in research.

In summary, our unified method efficiently links predictive accuracy, explainability, and causal reasoning key elements for utilizing trustworthy AI in clinical practice.

6. Conclusions

This study developed an explainable machine learning framework that is based on XGBoost for predicting heart disease and integrate SHAP explanation and causal validation using PC algorithm. As compared to traditional models, this framework not only carries correct predictions also explains which features truly contribute to the outcome this makes it more reliable for clinical decision-making. We also developed a user-friendly interface that allows clinicians to search predictions and explanations based on their expertise level.

Our research is the extension of Hou et al. [4] research, who applied causal discovery to SHAP outputs. Our framework, contrasting Hou et al. [4] work focuses on heart disease prediction and provides a clinical interface. The main difference between these two research our research concentrating precisely on cardiovascular risk prediction and it offers a real-time interaction and supports explanation customization making it more appropriate for real-world medical use.

On the UCI Heart Disease dataset, the model achieved high predictive accuracy and provided clinically meaningful insights. By adding causal analysis, we were able to confirm which features have real influence, reducing the risk of relying on misleading correlations.

Overall, our work advances explainable AI in healthcare by making predictions more interpretable, reliable, and accessible to medical professionals. Future work will expand dataset size (time-based clinical records), and increased field testing with clinicians in real-world environments.

References

1. World Health Organization. Cardiovascular diseases (CVDs). 2021. Available online: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) (accessed on 7 October 2025).
2. Rajkomar, A.; Dean, J.; Kohane, I. Machine learning in medicine. *N. Engl. J. Med.* 2019, *380*(14), 1347–1358.
3. Skouteli, C.; Prenzas, N.; Kakas, A.; Pattichis, C.S. Explainable AI modeling in the prediction of cardiovascular disease risk. In *Digital Health and Informatics Innovations for Sustainable Health Care Systems*; IOS Press: Amsterdam, the Netherlands, 2024; pp. 978–982.
4. Hou, Y.; Rabbani, S.B.; Hong, L.; Diawara, N.; Samad, M.D. Causal explainability of machine learning in heart failure prediction from electronic health records. *arXiv* 2025, arXiv:2506.03068.
5. Chen, T.; Guestrin, C. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
6. Yang, Y.; Liao, C.Y.; Keyvanshokooh, E.; Shao, H.; Weber, M.B.; Pasquel, F.J.; Garcia, G.G. A responsible framework for assessing, selecting, and explaining machine learning models in cardiovascular disease outcomes among people with type 2 diabetes: Methodology and validation study. *JMIR Med. Inform.* 2025, *13*, e66200.
7. Ganie, S.M.; Pramanik, P.K.; Zhao, Z. Ensemble learning with explainable AI for improved heart disease prediction based on multiple datasets. *Sci. Rep.* 2025, *15*, 13912.
8. Mienye, I.D.; Jere, N. Optimized ensemble learning approach with explainable AI for improved heart disease prediction. *Information* 2024, *15*(7), 394.
9. Moore, A.; Bell, M. XGBoost, a novel explainable AI technique, in the prediction of myocardial infarction: A UK Biobank cohort study. *Clin. Med. Insights Cardiol.* 2022, *16*, 11795468221133611.
10. Moore, A.; Bell, M. XGBoost, a novel explainable AI technique, in the prediction of myocardial infarction: A UK Biobank cohort study. *Clin. Med. Insights Cardiol.* 2022, *16*, 11795468221133611.
11. Tiwari, A.; Chugh, A.; Sharma, A. Ensemble framework for cardiovascular disease prediction. *Comput. Biol. Med.* 2022, *146*, 105624.
12. Asif, D.; Bibi, M.; Arif, M.S.; Mukheimer, A. Enhancing heart disease prediction through ensemble learning techniques with hyperparameter optimization. *Algorithms* 2023, *16*(6), 308.
13. Qi, X.; Wang, S.; Fang, C.; Jia, J.; Lin, L.; Yuan, T. Machine learning and SHAP value interpretation for predicting comorbidity of cardiovascular disease and cancer with dietary antioxidants. *Redox Biol.* 2025, *79*, 103470.
14. Tarabanis, C.; Kalampokis, E.; Khalil, M.; Alviar, C.L.; Chinitz, L.A.; Jankelson, L. Explainable SHAP-XGBoost models for in-hospital mortality after myocardial infarction. *Cardiovasc. Digit. Health J.* 2023, *4*(4), 126–132.
15. Gambetti, A.; Han, Q.; Shen, H.; Soares, C. A survey on human-centered evaluation of explainable AI methods in clinical decision support systems. *arXiv* 2025, arXiv:2502.09849.
16. Liu, Q.; Wong, W.H. An AI-powered Bayesian generative modeling approach for causal inference in observational studies. *arXiv* 2025, arXiv:2501.00755.
17. Lu, M.; Covert, I.; White, N.J.; Lee, S.I. CODE-XAI: Construing and deciphering treatment effects via explainable AI using real-world data. *medRxiv* 2024, 2024-09.
18. Perrett, A. Emergent explainability: Adding a causal chain to neural network inference. *arXiv* 2024, arXiv:2401.15840.
19. Almomani, O., Alsaaidah, A. ., Abu-Shareha, A. A. ., Alzaqebah, A. ., Amin Almaiah, M. ., & Shambour, Q. (2025). Enhance URL Defacement Attack Detection Using Particle Swarm Optimization and Machine Learning. Journal of Computational and Cognitive Engineering, 4(3), 296-308. https://doi.org/10.47852/bonviewJCCE52024668
20. Shmueli, G.; Martens, D.; Yoo, J.; Greene, T. From what-ifs to insights: Counterfactuals in causal inference vs. explainable AI. *arXiv* 2025, arXiv:2505.13324.
21. UCI Machine Learning Repository. Heart Disease Dataset. University of California, Irvine. Available online: https://archive.ics.uci.edu/ml/datasets/heart+Disease (accessed on 7 October 2025).
22. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-Sampling Technique. J. Artif. Intell. Res. 2002, 16, 321–357.
23. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.

24. Shambour, Qusai, Mahran Al-Zyoud, and Omar Almomani. "Quantum-Inspired Hybrid Metaheuristic Feature Selection with SHAP for Optimized and Explainable Spam Detection." Symmetry 17, no. 10 (2025): 1716.https://doi.org/10.3390/sym17101716
25. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In *Advances in Neural Information Processing Systems (NeurIPS 2017)*, Long Beach, CA, USA, 4–9 December 2017; *30*.
26. Glymour, C.; Zhang, K.; Spirtes, P. Review of Causal Discovery Methods Based on Graphical Models. *Front. Genet.* 2020, *11*, 524.
27. Hussain, M., Chen, C., Hussain, M. et al. Optimised knowledge distillation for efficient social media emotion recognition using DistilBERT and ALBERT. Sci Rep 15, 30104 (2025). https://doi.org/10.1038/s41598-025-16001-9
28. Streamlit Inc. Streamlit: The Fastest Way to Build and Share Data Apps. Available online: https://streamlit.io/ (accessed on 28 October 2025).