

# Explainable Machine Learning for Identifying False-Negative Myocardial Infarction: A SHAP-Based Analysis

Mathesh Anandan, Shirly Sasikumar, Darius Alan Arockiasamy, Aadhithan Dhanasekaran, Joanna Phebe Augustine, Dhivya Kothandan\*

Department of Pharmacy Practice, C.L. Baid Metha College of Pharmacy, Thoraipakkam, Chennai, Tamil Nadu, INDIA.

## ABSTRACT

**Background:** The timely and accurate identification of Myocardial Infarction (MI) remains a challenge in clinical practice. Diagnostic inconsistency may result in missed or delayed diagnoses, especially in resource-limited settings. This study aimed to develop an explainable Machine Learning (ML) based framework to detect MI cases that may be overlooked by conventional biomarker-based diagnosis, particularly in patients with normal troponin and Creatine Kinase-MB (CK-MB) levels, by using SHapley Additive exPlanations (SHAP). **Materials and Methods:** A retrospective dataset comprising 1,319 patient records with nine clinical variables was utilized. Multiple ML classifiers, including Random Forest, XGBoost, and Decision Tree, were trained following pre-processing steps: normalization, feature encoding, hyperparameter tuning. Model performance was assessed using accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). SHAP analysis was employed to elucidate feature contributions and support model predictions. **Results:** Tree-based classifiers achieved high predictive performance, with accuracies exceeding 97% and AUC values up to 0.98. SHAP analysis identified troponin, CK-MB, and age as the most influential predictors of myocardial infarction. Importantly, the model correctly identified MI in patients with normal troponin and CK-MB levels, by integrating biomarker patterns with demographic and hemodynamic features. These findings indicate that the explainable ML framework can detect MI cases at risk of false-negative classification using conventional biomarker thresholds. **Conclusion:** The integration of ML with explainable Artificial Intelligence (AI) techniques shows strong potential for early and accurate MI prediction, including the identification of cases that may be falsely classified as negative using conventional diagnostic approaches. This approach supports improved clinical decision making, enhances diagnostic safety by reducing false-negative outcomes.

**Keywords:** CK-MB, Myocardial infarction, Machine learning, Risk prediction, SHAP, Troponin.

## Correspondence:

**Dr. Dhivya Kothandan**

Associate Professor, Department of Pharmacy Practice, C.L. Baid Metha College of Pharmacy, Thoraipakkam, Chennai-600097, Tamil Nadu, INDIA.  
Email: divyapharmd@gmail.com

**Received:** 24-12-2025;

**Revised:** 06-03-2026;

**Accepted:** 15-05-2026.

## INTRODUCTION

Cardiovascular Diseases (CVDs) are the leading cause of death worldwide, about 17.9 million deaths annually (32% of total deaths), with Myocardial Infarction (MI) and stroke nearly 85%. In India, CVDs account for 28.1% of deaths, with ischemic heart diseases and strokes representing nearly 80% (Lalit, 2018). Diagnosis of MI relies on a combination of clinical history, Electrocardiogram (ECG) findings, and biomarkers indicative of myocardial injury. Cardiac troponins, specifically Troponin T (cTnT) and Troponin I (cTnI), are considered gold-standard

markers due to their high sensitivity and specificity for myocardial necrosis. Imaging studies, such as chest X-rays are often used to rule out alternative causes of chest pain (Newby *et al.*, 2012).

Advancements in diagnostic technologies have contributed significantly to the MI prediction and detection, with Machine Learning (ML) algorithms, identifying hidden patterns in patient data for early risk prediction, though their clinical efficacy remains debated (Khera *et al.*, 2021). Artificial Intelligence (AI) based deep learning models analyze ECGs not only to detect MI, but also arrhythmias, and heart failure. Additionally, Magnetic Field Imaging (MFI) or Magnetocardiography (MCG) provides non-invasive detection of cardiac signals, aiding in early MI diagnosis (Schlapfer *et al.*, 2017). Wearable bio signal sensors, such as wrist-worn transdermal devices, allow for the rapid detection of elevated cTnI levels without the need for blood draws, facilitating real-time monitoring for prompt MI detection (Sengupta *et al.*, 2023).



DOI: 10.5530/jyp.20260004

### Copyright Information :

Copyright Author (s) 2026 Distributed under Creative Commons CC-BY 4.0

Publishing Partner : Manuscript Technomedia. [www.mstechnomedia.com]

Despite advancements in diagnostic technologies, the timely and accurate identification of MI remains a critical challenge in clinical practice. Diagnostic inconsistency and dependence on physician interpretation may result in missed or delayed diagnoses, especially in resource-limited settings. These limitations highlight the need for data-driven, and interpretable tools that can enhance clinical decision-making. ML has the potential to analyze complex patterns in large datasets and improve predictive accuracy beyond conventional methods. However, the “black-box” nature of many ML models poses challenges for clinical adoption (Chowdhury *et al.*, 2025).

Therefore, integrating explainable AI method such as SHapley Additive exPlanations (SHAP) enhances model transparency and strengthens clinician trust in ML based predictions. This is particularly important for detecting MI cases that may be overlooked by conventional diagnostic methods, thereby improving patient safety and outcomes. In this study, a ML model was developed to detect patients at risk of false-negative MI by considering subtle clinical features and biomarker patterns beyond standard diagnostic thresholds.

## MATERIALS AND METHODS

### Study Design and Data Source

This study was conducted as a secondary data analysis using a publicly available dataset obtained from the Mendeleev Data repository (Maghdid *et al.*, 2022). This dataset contains information on 1,319 patients across 9 variables. Of these, 8 variables, including age, gender, heart rate, Systolic Blood Pressure (SBP), Diastolic BP (DBP), blood sugar, CK-MB, and troponin, were used as predictors. The last variable (Output) indicates the presence or absence of MI. A supervised ML approach was employed to develop a predictive model for detecting false-negative MI. This study also incorporated SHAP to enhance the interpretability of model predictions.

### Classifier Algorithm

Classifiers are a set of ML algorithms designed for classification tasks, such as predicting whether the patient has the disease or not. In this study, we used several classifiers to train our model, including the decision tree classifier, random forest classifier, Support Vector Classifier (SVC), logistic regression, naive Bayes classifier, gradient boost classifier, XGBoost classifier, LightGBM classifier, and CatBoost classifier.

### Data Pre-processing

The dataset was examined for completeness prior to analysis, and no missing values were identified. Therefore, data imputation procedures were not required. Before training the dataset with ML algorithms, detailed data exploration and preprocessing were carried out to ensure the dataset was suitable for analysis. Although several outliers were detected, they were retained

due to the limited size of the dataset. Also, most of the selected classifiers can handle outliers effectively. The gender feature was encoded as 1 representing male and 0 representing female. All numerical features were scaled to a range between 0 and 1 using the MinMaxScaler method to maintain uniformity and improve model performance. Normalization reduces bias in model interpretation and provides more accurate predictions. To minimize the risk of data leakage, normalization parameters were derived exclusively from the training dataset and subsequently applied to the testing data. Additionally, hyperparameter tuning was performed using grid search combined with k-fold cross-validation to reduce the risk of model overfitting and improve generalization.

The formula for Min-Max scaling is,

$$X_{\text{scaled}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}$$

Where X represents the original value,  $X_{\text{min}}$  is the minimum value of the feature,  $X_{\text{max}}$  is the maximum value, and  $X_{\text{scaled}}$  is the normalized value between 0 and 1.

### Model Training and Evaluation

The pre-processed data was randomly divided into training and testing subsets in an 80:20 ratio. The training set was utilized to train various classification algorithms. To improve model performance and ensure reliable results, hyperparameter tuning was performed using grid search combined with cross-validation. After training, the model was evaluated with test data. The dataset consisted of 810 MI cases and 509 non-MI cases, representing a moderately balanced distribution. Performance metrics such as precision, recall, and F1-score were therefore evaluated alongside accuracy and Area Under the Curve (AUC) to provide a more comprehensive assessment of classification performance.

### SHAP Interpretation

SHAP is a widely adopted framework for interpreting ML models. It explains the contribution of each feature to an individual prediction using Shapley values from game theory, ensuring fair attribution. This improves transparency by revealing how different variables influence the model's decision. Such interpretability is essential in healthcare, where clinical decisions must be accurate and trustworthy. SHAP values typically explain how each variable contributes to the model's output. Initially, it estimates the base value, which is the average model's output over an entire dataset. For a binary classification model, the predicted probability for an individual instance can be computed as:

$$P(\text{class } 1) = \text{Base Value} + \sum \text{SHAP values for all features}$$

Similarly, the probability for class 0 is:

$$P(\text{class } 0) = 1 - P(\text{class } 1)$$

## Statistical Analysis

All statistical analyses were performed using Python (version 3.8), utilizing statistical libraries such as NumPy, pandas, and SciPy. A statistical significance was indicated by a  $p$ -value of less than 0.05. Independent sample  $t$ -test was performed to compare the mean values of continuous variables between the two groups (patients diagnosed with MI and those without MI). One way Analysis of Variance (ANOVA) is used to find out the statistical difference between different age groups.

## RESULTS

### Exploratory Data Analysis

An exploratory data analysis was performed to understand the dataset thoroughly and to optimize model training tailored to its features. Table 1 summarizes the mean values of clinical features stratified by MI status. Patients with MI had a significantly higher mean age ( $58.77 \pm 12.96$  years) compared to non-MI patients ( $52.09 \pm 13.73$  years), with a test statistic of 8.78 and  $p < 0.05$ . Similarly, CK-MB levels were markedly elevated in the MI group ( $23.27 \pm 57.70$  ng/mL) relative to the non-MI group ( $2.56 \pm 1.37$  ng/mL), with a test statistic of 10.21 and  $p < 0.05$ . Troponin levels were also substantially higher in MI patients ( $0.57 \pm 1.39$  ng/mL) compared to non-MI individuals ( $0.03 \pm 0.44$  ng/mL), with a test statistic of 10.33 and  $p < 0.05$ , emphasizing their strong association with MI.

The dataset comprises 810 patients diagnosed with MI and 509 without MI, including 870 males and 449 females. As illustrated in Figure 1, age is a major contributing factor to MI. The incidence of MI increases markedly with advancing age, particularly beyond 40 years, where a considerable rise in MI cases is observed compared to non-MI cases.

Table 2 presents the age-wise analysis of clinical features. Troponin and CK-MB levels, both significant markers for MI, show a clear upward trend with increasing age. Troponin levels begin to rise

from the 40-50 age group ( $0.31 \pm 0.90$  ng/mL), continuing to increase in the 50-60 ( $0.40 \pm 1.13$  ng/mL) and 60-70 ( $0.49 \pm 1.48$  ng/mL) age groups with  $p$ -value of  $< 0.05$ , indicating a higher likelihood of myocardial injury with age. Similarly, CK-MB levels, which are more variable, show elevated mean values in the 60-70 ( $17.49 \pm 52.31$  ng/mL) and 80-90 ( $32.58 \pm 70.96$  ng/mL) age groups.

### ML Model Performance

The performance metrics of all models are shown in Table 3. Ensemble models, including Random Forest, XGBoost, LightGBM, Gradient Boosting, and CatBoost, achieved the highest performance in predicting MI, with 98% accuracy along with strong precision, recall, F1-score, and AUC across both MI and non-MI classes. The decision tree classifier also performed well, attaining an accuracy of 97.5%. Although Naive Bayes exhibited less accuracy (88.7%), it achieved a high precision for the MI group, making it useful for minimizing false negatives. Logistic regression (76.3% accuracy) and SVC (67.2% accuracy, 30% MI recall) underperformed, with SVC showing the highest rate of misclassification risk. Overall, tree-based ensemble models proved to be the most effective and reliable for MI prediction.

The performance of the classification models was evaluated using confusion matrices, which reveal the distribution of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), as summarized in Table 4. Ensemble models such as Random Forest, XGBoost, and LightGBM showed strong performance, achieving high TP rates around 37% and TN rates near 60%, while keeping FP and FN rates relatively low. Naive Bayes also performed competitively, with a TP rate of 38.18%, although it exhibited a higher FN rate. In contrast, the SVC underperformed with a TP of only 11.52% and a high FP rate of 27.27%, making it less reliable for MI classification.

**Table 1: Comparison of Features Between Patients with MI and without MI.**

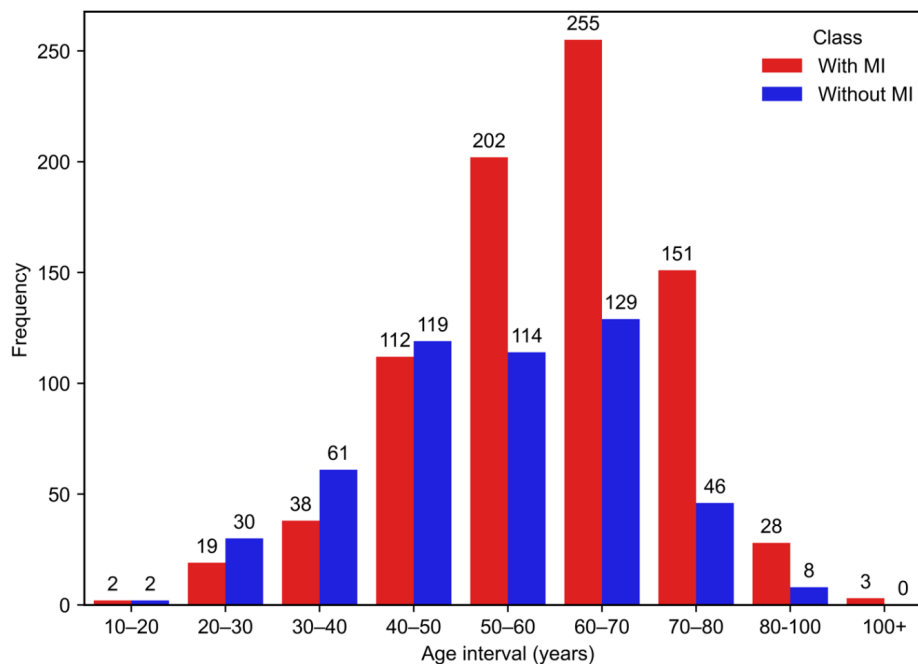
Sl. No.	Feature	Total	Without MI	With MI	Test Statistics	$p$ value
1	Age (years)	$56.19 \pm 13.65$	$52.09 \pm 13.73$	$58.77 \pm 12.96$	8.78	0.0321*
2	Heart Rate (bpm)	$78.34 \pm 51.63$	$77.89 \pm 48.21$	$78.62 \pm 53.69$	0.24	0.812
3	SBP (mmHg)	$127.17 \pm 26.12$	$127.86 \pm 27.04$	$126.74 \pm 25.54$	-0.75	0.455
4	DBP (mmHg)	$72.27 \pm 14.03$	$72.44 \pm 14.33$	$72.16 \pm 13.86$	-0.35	0.727
5	Glucose (mg/dL)	$146.63 \pm 74.92$	$149.76 \pm 78.41$	$144.67 \pm 72.63$	-1.18	0.238
6	CK-MB (ng/mL)	$15.27 \pm 46.33$	$2.56 \pm 1.37$	$23.27 \pm 57.70$	10.21	0.0316*
7	Troponin (ng/mL)	$0.36 \pm 1.15$	$0.03 \pm 0.44$	$0.57 \pm 1.39$	10.33	0.0223*

Values are expressed as Mean  $\pm$  Standard Deviation (SD). Comparisons between groups were performed using the independent samples  $t$ -test. SBP: systolic blood pressure; DBP: diastolic blood pressure; CK-MB: creatine kinase-MB. A  $p$  value  $< 0.05$  was considered statistically significant.

**Table 2: Features Based on Age Group.**

Sl. No.	Age Group (years)	Troponin (ng/mL)	CK-MB (ng/mL)	Systolic BP (mmHg)	Diastolic BP (mmHg)	Glucose (mg/dL)	Heart Rate (bpm)
1	10-20	0.31±0.74	57.96±119.33	120.5±17.64	67.5±7.55	160.0±85.80	71.16±10.81
2	20-30	0.07±0.29	19.67±58.03	129.42±33.39	71.43±17.14	147.82±86.37	77.26±15.09
3	30-40	0.06±0.26	18.04±52.86	124.39±26.26	71.08±12.74	142.91±62.84	83.39±94.49
4	40-50	0.31±0.90	12.31±34.83	126.70±26.53	72.71±14.83	143.35±69.70	80.13±61.59
5	50-60	0.40±1.13	12.86±38.94	127.30±25.50	72.57±13.49	153.77±79.52	75.6±14.34
6	60-70	0.49±1.48	17.49±52.31	128.00±26.44	72.51±14.11	147.30±80.28	78.67±51.11
7	70-80	0.45±1.26	12.91±45.32	126.29±24.58	71.65±14.06	142.0±71.14	81.84±73.48
8	80-90	0.30±0.73	32.58±70.96	127.42±23.77	71.82±14.29	141.25±73.21	74.17±16.82
9	90-100	0.53±0.79	5.65±0.37	135.8±53.93	66.4±9.68	223.2±145.13	76.4±28.65
10	100-110	0.31±0.45	42.31±32.90	175.33±31.72	74.33±18.00	199.66±113.00	67.33±16.28
	p-value	0.0200*	0.0658	0.1389	0.9277	0.2142	0.9233

Values are expressed as mean±Standard Deviation (SD). Troponin and CK-MB concentrations are reported in ng/mL. SBP: systolic blood pressure; DBP: diastolic blood pressure. Comparisons across age groups were performed using one-way analysis of variance (ANOVA). A *p* value <0.05 was considered statistically significant.



**Figure 1:** Number of Patients with and without MI. Red bars represent patients with MI, and blue bars represent patients without MI. Values displayed above each bar indicate the number of subjects in the corresponding age group. The x-axis denotes age intervals (years), and the y-axis represents frequency (Reproduce at column width).

## SHAP Interpretation

SHAP TreeExplainer was used to interpret the predictions of the Random Forest model for MI. The SHAP summary plot (Figure 2) ranks feature by their impact on the model's output. Among the features, troponin and CK-MB showed the highest impact, with elevated values (indicated in red) significantly increasing the predicted probability of MI and lower values (in blue) reducing it. Age also contributed positively to MI risk, with higher age values

correlating with increased prediction scores. Blood pressure showed a moderate influence, whereas glucose and heart rate demonstrated lower and more variable effects across the dataset.

The SHAP Waterfall plot (Figure 3) depicts the model's correct prediction of MI for a patient with a high prediction score of 0.78. Despite troponin being within the normal range (0.016 ng/mL), it still contributes positively to the prediction along, indicating that the model has learned to identify subtle patterns associated

with MI risk and SBP also has a minimal impact toward positive direction. Protective factors like CK-MB (2.78 ng/mL) and age (36 years) slightly reduce the overall risk but are not sufficient to alter the final outcome. This figure highlights the model's ability to capture complex interactions among variables, even when primary biomarkers appear normal.

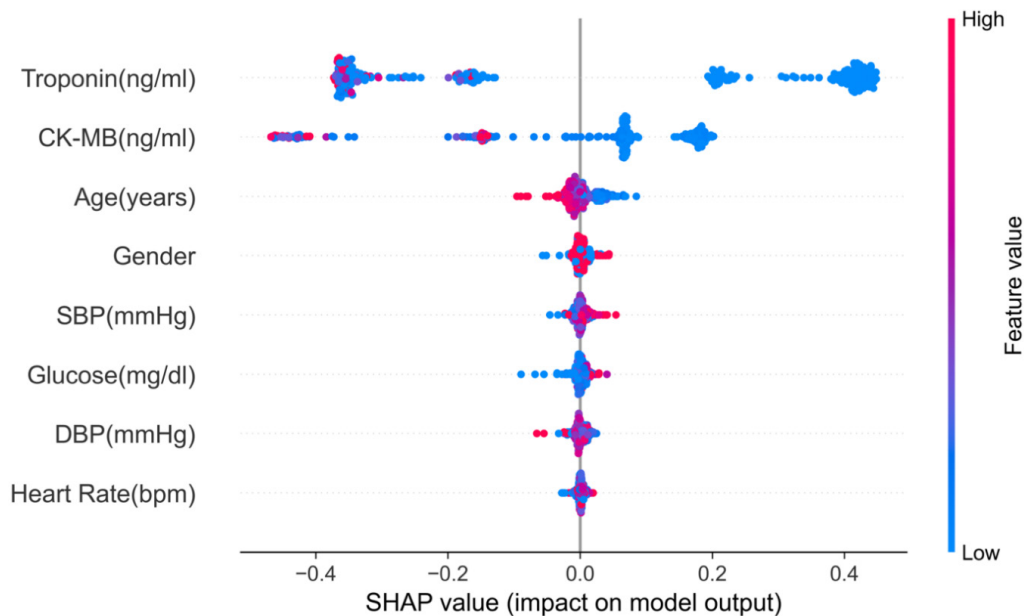
Figure 4 presents the model's correct prediction of MI for a patient with a high risk score of 0.82. CK-MB (8.87 ng/mL) falls

within the normal range, it exerts a strong positive influence on the prediction, suggesting that the model detects subtle risk patterns. Troponin (0.005 ng/mL), heart rate (119 bpm), Glucose (184 mg/dL) and Gender (Male encoded as 1) slightly lower the risk, but not sufficiently to offset the impact of CK-MB. Although cardiac biomarkers remain the strongest predictors of myocardial infarction, the model demonstrated the ability to integrate additional clinical variables such as age, heart rate, glucose levels, and blood pressure to identify MI cases even when biomarker

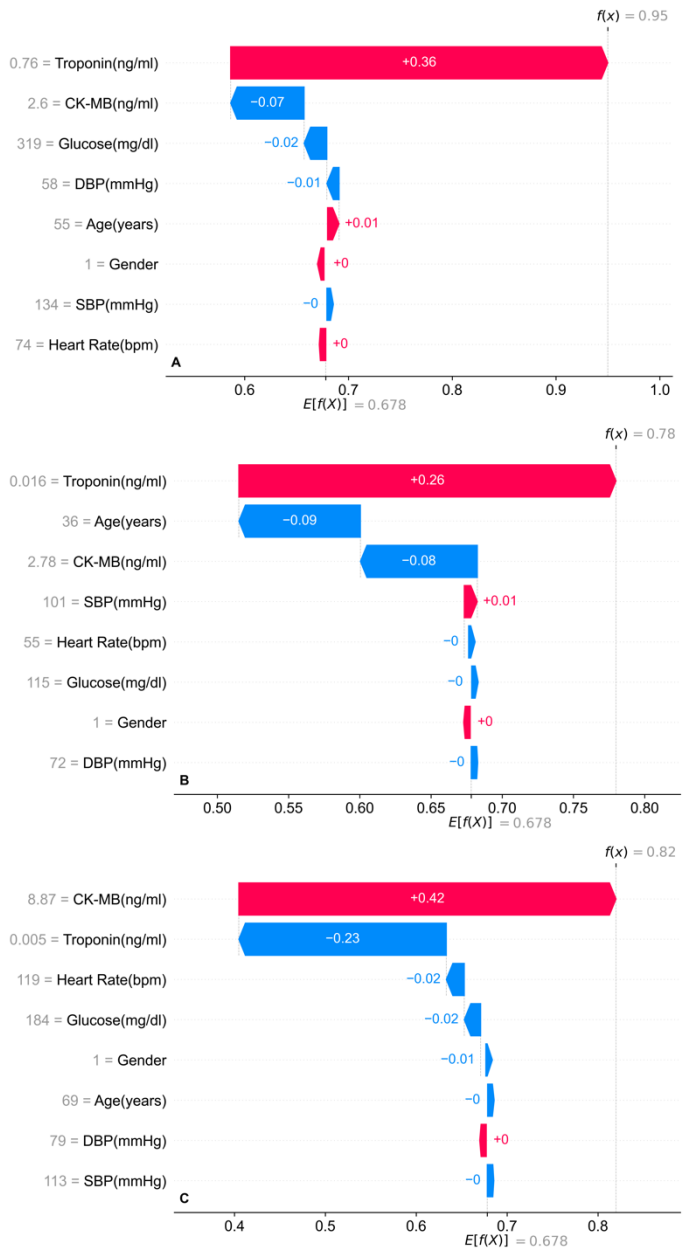
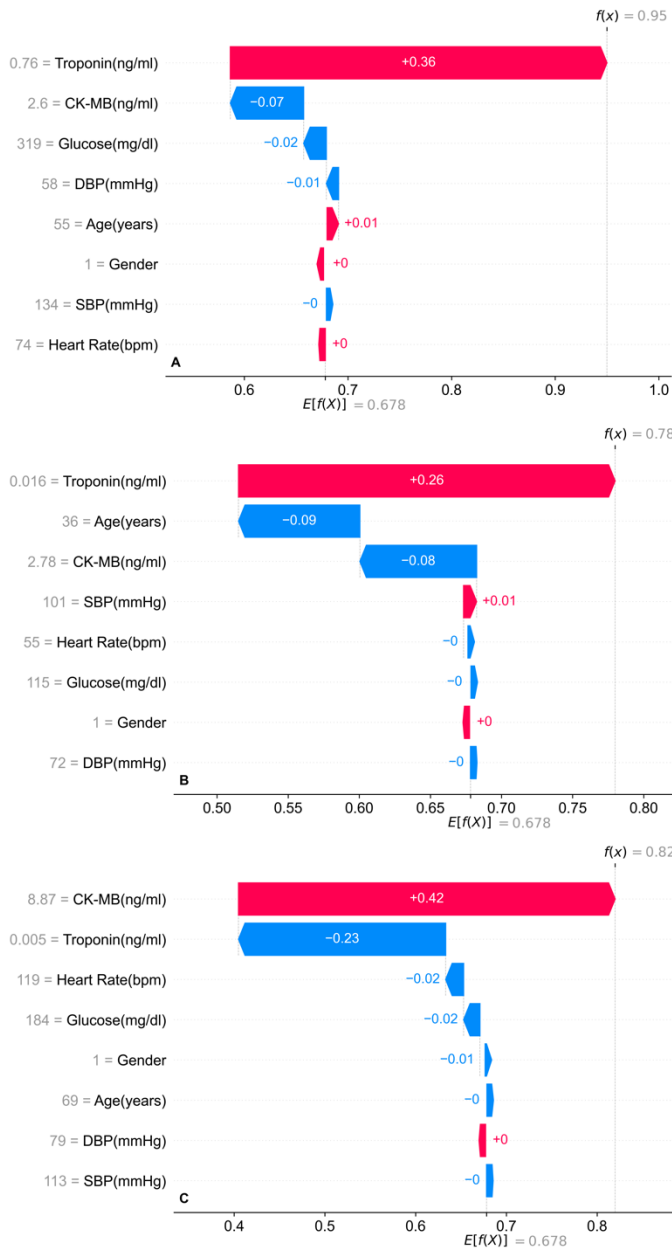
**Table 3: Performance of Models.**

Sl. No.	Model	Accuracy (%)	Precision (%)		Recall (%)		F-1 score (%)		AUC (%)
			Without MI	With MI	Without MI	With MI	Without MI	With MI	
1	Decision Tree Classifier	97.5	98	98	96	99	97	98	97
2	Random Forest Classifier	98	98	98	97	99	97	98	98
3	SVC	67.2	68	67	30	91	41	77	60
4	Logistic Regression	76.3	71	80	66	83	69	81	74
5	Navie Bayes	88.7	78	99	98	83	87	90	90
6	Gradient Boosting	97.9	98	97	95	99	97	98	98
7	XGBoost	98	98	98	96	99	97	98	98
8	LightBM	98.1	98	98	96	99	97	98	98
9	CatBoost	98	97	97	97	98	97	98	98

SVC: Support Vector Classifier. Model performance is reported in terms of accuracy, precision, recall, F1-score, and Area Under the Curve (AUC). Precision, recall, and F1-score are presented separately for patients without MI and with MI. All values are expressed as percentages (%).



**Figure 2:** SHAP Summary Plot for MI Prediction. Each point represents an individual observation. The x-axis shows the SHAP value, indicating the impact of each feature on the model output, with positive values increasing and negative values decreasing the predicted probability of MI. Features are ordered by mean absolute SHAP value. The color scale represents the feature value, with red indicating higher values and blue indicating lower values (Reproduce at column width).



**Figure 3:** SHAP Waterfall plot for Individual Patient with normal Troponin and experienced MI. Troponin and systolic blood pressure contribute positively despite normal biomarker values, while age and CK-MB show minor protective effects (Reproduced at column width).

**Figure 4:** SHAP Waterfall plot for Individual Patient with normal CK-MB and experienced MI. CK-MB is the dominant contributor to the prediction, with other features exerting minimal opposing influence. Red bars indicate positive and blue bars indicate negative contributions to the model output (Reproduced at column width).

values fall within normal diagnostic thresholds. This capability highlights the potential role of ML-based systems in assisting clinicians in detecting possible false-negative cases.

## DISCUSSION

MI is a major health concern, accounting for 31.7% of deaths in India and causing significant physical and psychological burden (Jan *et al.*, 2024). Despite advances in ECG and biomarker testing, some MI cases remain undetected, particularly when biomarker levels are inconclusive, leading to false-negative

diagnoses. ML can address this limitation by identifying complex patterns beyond fixed thresholds. In our study, CK-MB levels differed significantly between MI and non-MI groups (Table 1), consistent with (Shamshirian *et al.*, 2020) study. Age emerged as an important predictor, with most MI cases occurring in the 60-70 years group, highlighting the need for age-diverse datasets. Troponin levels also increased significantly with age ( $P = 0.002$ ), aligning with the study of (Liu *et al.*, 2017). These results emphasize the importance of age- and sex-specific biomarker interpretation and demonstrate the potential of ML to improve MI detection accuracy.

**Table 4: Confusion Matrix.**

Sl. No.	Model	True Positive (%)	True Negative (%)	False Positive (%)	False Negative (%)
1	Decision Tree Classifier	37.27	60.3	1.52	0.91
2	Random Forest Classifier	37.27	60.0	1.52	1.21
3	SVC	11.52	55.76	27.27	5.45
4	Logistic Regression	25.76	50.61	13.03	10.61
5	Navies Bayers	38.18	50.61	0.61	10.61
6	Gradient Boosting	36.97	60.61	1.82	0.61
7	XGBoost	37.27	60.61	1.52	0.61
8	LightGBM	37.27	60.3	1.52	0.91
9	CatBoost	37.58	60.0	1.21	1.21

SVC: Support Vector Classifier. True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) rates are expressed as percentages (%). These metrics summarize the classification outcomes of each model in distinguishing patients with and without MI.

Compared with the XGBoost model reported by (Moore *et al.*, 2022), which achieved 75% accuracy and an AUC of 0.86 with poor MI precision (0.07), our XGBoost model demonstrated markedly improved performance, achieving 98% accuracy, 0.98 AUC, and balanced precision-recall (Table 3). Similarly, while (Soleimani *et al.*, 2025) reported AUC values of 0.924 and 0.905 for Random Forest and XGBoost, respectively, our ensemble models (XGBoost, LightGBM, and CatBoost) achieved superior results, with 98% accuracy, AUC, precision, and recall. Consistent with prior findings, our SVC model showed comparatively lower performance. Furthermore (Vicente *et al.*, 2023), reported LightGBM accuracy of 92.6% and an F1-score of 0.922 for MI mortality prediction, whereas our LightGBM model achieved 98.1% accuracy and an F1-score of 0.98, reflecting the effectiveness of improved preprocessing and optimization strategies.

This study addresses concerns about overinterpreting feature-outcome relationships in ML models (Van Smeden *et al.*, 2022) by using SHAP to provide transparent, case-level explanations without implying causality. This approach improves interpretability while maintaining high predictive performance, overcoming a key limitation of black-box models in clinical decision-making. The importance of explainability is supported by (Diprose *et al.*, 2020), who reported that 88% of physicians preferred predictions accompanied by clear explanations, as they improved understanding and trust in model outputs.

This study has certain limitations. First, the analysis was based on a relatively small, single publicly available retrospective dataset, which may limit the generalizability of the findings. In addition, the dataset lacked lifestyle-related variables and detailed temporal information on biomarker measurements, both of which may influence diagnostic accuracy. Future studies should therefore include larger, multi-center datasets with more diverse populations, incorporate temporal biomarker

data, and consider additional clinical and lifestyle variables to improve model robustness and external validity. Further stratified analyses focusing on patients with normal biomarker levels would also help better evaluate the model's ability to detect potential false-negative MI cases. For clinical translation, ML-based prediction models could be integrated into hospital decision-support systems to assist clinicians during the early assessment of suspected MI. The use of explainable AI techniques such as SHAP may enhance clinician trust by clearly illustrating how individual features influence predictions. Nevertheless, prospective clinical validation and careful determination of optimal decision thresholds will be essential before routine clinical implementation.

## CONCLUSION

This study demonstrates the clinical value of machine learning models, particularly the Random Forest algorithm, in detecting MI cases that may be overlooked by conventional diagnostic methods. This is especially relevant when commonly used biomarkers such as troponin and CK-MB appear normal or inconclusive. SHAP analysis enhanced model transparency by clearly explaining how individual features influenced predictions, thereby improving clinician confidence. By integrating a wider range of clinical and physiological variables beyond standard biomarkers, the model achieved more accurate risk prediction, particularly in complex cases. Overall, these findings highlight the potential of AI-based decision support systems to facilitate earlier diagnosis, support informed clinical decisions, and improve patient outcomes in cardiovascular care.

## ACKNOWLEDGEMENT

The authors express their gratitude to the Mendeley Data for providing access to the MI dataset and acknowledge the dataset contributors for making this valuable resource openly available.

## ABBREVIATIONS

**AI:** Artificial Intelligence; **AMI:** Acute Myocardial Infarction; **ANOVA:** Analysis of Variance; **AUC:** Area Under the Curve; **CABG:** Coronary Artery Bypass Grafting; **CK-MB:** Creatine Kinase-Myocardial Band; **cTnI:** Cardiac Troponin I; **cTnT:** Cardiac Troponin T; **CVD:** Cardiovascular Disease; **CVDs:** Cardiovascular Diseases; **DBP:** Diastolic Blood Pressure; **ECG:** Electrocardiogram; **FN:** False Negative; **FP:** False Positive; **GBM:** Gradient Boosting Machine; **hs-cTnT:** High-Sensitivity Cardiac Troponin T; **LightGBM:** Light Gradient Boosting Machine; **MCG:** Magnetocardiography; **MFI:** Magnetic Field Imaging; **MI:** Myocardial Infarction; **ML:** Machine Learning; **NSTEMI:** Non-ST-Elevation Myocardial Infarction; **PCI:** Percutaneous Coronary Intervention; **SBP:** Systolic Blood Pressure; **SHAP:** SHapley Additive exPlanations; **SVC:** Support Vector Classifier; **SVM:** Support Vector Machine; **STEMI:** ST-Elevation Myocardial Infarction; **T2MI:** Type 2 Myocardial Infarction; **TN:** True Negative; **TP:** True Positive; **XGBoost:** Extreme Gradient Boosting.

## CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

## REFERENCES

- Chowdhury, M.A., Rizk, R., Chiu, C., Zhang, J.J., Scholl, J.L., Bosch, T.J., Singh, A., Baugh, L.A., McGough, J.S., Santosh, K., et al. (2025). The Heart of Transformation: Exploring Artificial Intelligence in Cardiovascular Disease. *Biomedicine*, 13(2), 427. <https://doi.org/10.3390/biomedicine13020427>
- Jan., Bisma., Dar., Mohammad Imran., Choudhary, Bharti., Basist., Parakh., Khan., Rahmuddin., Alhalmi., Abdulsalam. (2024). Cardiovascular Diseases Among Indian Older Adults: A Comprehensive Review. *Cardiovascular therapeutics*, 2024, 6894693. <https://doi.org/10.1155/2024/6894693>
- Jie-Yin Liu., Qiao-Wei Jia., Xiao-Ling Zang., Rong-Hu Wang., Chun-Jian Li., Lian-Sheng Wang., Wen-Zhu Ma., Zhi-Jian Yang., En-Zhi Jia. (2017). Age-sex distribution of patients with high-sensitivity troponin T levels below the 99th percentile. *Oncotarget*, 8(43), 75638-75645. <https://doi.org/10.18632/oncotarget.20328>
- Khera R., Haimovich J., Hurley NC., et al. (2021). Use of Machine Learning Models to Predict Death After Acute Myocardial Infarction. *JAMA cardiology*, 6(6), 633-641. <http://doi.org/10.1001/jamacardio.2021.0122>
- Lalit Dandona., India State-Level Disease Burden Initiative CVD Collaborators (2018). The changing patterns of cardiovascular diseases and their risk factors in the states of India: the Global Burden of Disease Study 1990-2016. *The Lancet. Global health*, 6(12), e1339-e1351. [https://doi.org/10.1016/S2214-109X\(18\)30407-8](https://doi.org/10.1016/S2214-109X(18)30407-8)
- Maarten van Smeden., Georg Heinze., Ben Van Calster., Folkert W Asselbergs., Panos E Vardas., Nico Bruining., Peter de Jaegere., Jason H Moore., Spiros Denaxas., Anne Laure Boulesteix., Karel G M Moons. (2022). Critical appraisal of artificial intelligence-based prediction models for cardiovascular disease. *European heart journal*, 43(31), 2921-2930. <https://doi.org/10.1093/eurheartj/ehac238>
- Maghdid S., Rashid T. An extensive dataset for the heart disease classification system. *Mendeley Data V1*. 2022. <https://doi.org/10.17632/65gxyg2nmg.1>
- Moore, A., Bell, M. (2022). XGBoost, A Novel Explainable AI Technique, in the Prediction of Myocardial Infarction: A UK Biobank Cohort Study. *Clinical Medicine Insights. Cardiology*, 16, 11795468221133611. <https://doi.org/10.1177/11795468221133611>
- Newby, L., Jesse, R., Babb, J., et al. (2012). ACCF 2012 expert consensus document on practical clinical considerations in the interpretation of troponin elevation: a report of the American College of Cardiology Foundation task force on Clinical Expert Consensus Documents. *Journal of the American College of Cardiology*, 60(23), 2427-2463. <https://doi.org/10.1016/j.jacc.2012.08.969>
- Schlapfer, J., Wellens, H. J. (2017). Computer-interpreted electrocardiograms: Benefits and limitations. *Journal of the American College of Cardiology*, 70(9), 1183-1192. <https://doi.org/10.1016/j.jacc.2017.07.723>
- Shamshirian, A., Alizadeh-Navaei, R., Abedi, S., Jafarpour, H., Fazli, H., Hosseini, S., Hessami, A., Karimifard, K., Yosefi, S., Zahedi, M., Motamen, S., Ghorbanpour, A., Zarandi, B., Esfahani, A., Rostamian-Moghaddam, Y., Mehdi-pour, S., Heydari, K., Aghajanian, S., Pour Mehdi, S., Mehdi, P., Azizi, S. (2020). Levels of Blood Biomarkers among Patients with Myocardial Infarction in Comparison to Control Group. *Ethiopian journal of health sciences*, 30(1), 5-12. <https://doi.org/10.4314/ejhs.v30i1.2>
- Shantanu Sengupta., Siddharth Biswal., Jitto Titus., Atandra Burman., Keshav Reddy., Mahesh C Fulwani., Aziz Khan., Niteen Deshpande., Smit Shrivastava., Naveena Yanamala., Partho P Sengupta. (2023). A novel breakthrough in wrist-worn transdermal troponin-I-sensor assessment for acute myocardial infarction. *European heart journal. Digital health*, 4(3), 145-154. <https://doi.org/10.1093/ehjdh/ztad015>
- Soleimani, H., Najdaghi, S., Davani, D.N., Dastjerdi, P., Samimisedeh, P., Shayesteh, H., Sattartabar, B., Masoudkabar, F., Ashraf, H., Mehrani, M., Jenab, Y. and Hosseini, K. (2025). Predicting In-Hospital Mortality in Patients With Acute Myocardial Infarction: A Comparison of Machine Learning Approaches. *Clinical cardiology*, 48(4), e70124. <https://doi.org/10.1002/clc.70124>
- Vicente, Ana Leticia Garcez., Roseval Donisete Malaquias Junior., and Roseli AF Romero. (2023). Explainable lightgbm approach for predicting myocardial infarction mortality. In *2023 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 323-328). IEEE. <https://doi.org/10.48550/arXiv.2404.15029>
- William K Diprose., Nicholas Buist., Ning Hua., Quentin Thurier., George Shand., Reece Robinson. (2020). Physician understanding, explainability, and trust in a hypothetical machine learning risk calculator. *Journal of the American Medical Informatics Association: JAMIA*, 27(4), 592-600. <https://doi.org/10.1093/jamia/oc229>

**Cite this article:** Mathesh A, Shirly S, Alan DA, Aadithan DS, Phebe J, Dhivya K. Explainable Machine Learning for Identifying False-Negative Myocardial Infarction: A SHAP-Based Analysis. *J Young Pharm.* 2026;18(2):473-80.