ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206

Volume 9, Issue 4, April 2026

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Explainable AI-Based Heart Disease Prediction and Clinical Insight System

Khan Jarjish[1], Shaikh Abdullah Muhammad Aslam[2], Huzaifa Saiyad[3], Khokar Saad Mohammad Rais[4], Prof. Ramkumar Maurya[5], Prof. Junaid Mandviwala[6], Dr. Varsha Shah[7]

UG Students, Dept. of AI&DS, Rizvi College of Engineering, Mumbai, Maharashtra, India[1,2,3,4]

Project Guide, Head of Department, Dept. of AI&DS, Rizvi College of Engineering, Mumbai, Maharashtra, India[5]

Dean of Academics, Rizvi College of Engineering, Mumbai, Maharashtra, India[6]

Principal, Rizvi College of Engineering, Mumbai, Maharashtra, India[7]

ABSTRACT: This project develops an intelligent system to predict heart disease using machine learning models like Logistic Regression, Decision Trees, Random Forest, and SVM. It uses patient health data (age, blood pressure, cholesterol, etc.) and evaluates models using metrics such as accuracy and precision.

A key feature is the use of Explainable AI techniques like SHAP and LIME, which provide clear insights into how predictions are made. The system includes a user-friendly interface that shows predictions along with visual explanations, helping clinicians understand risk factors.

Overall, the system improves diagnostic accuracy, supports clinical decision-making, and promotes transparent, data-driven healthcare.

KEYWORDS: Explainable AI, Heart Disease Prediction, Machine Learning, SHAP, LIME, Clinical Decision Support

I. INTRODUCTION

Heart disease is one of the leading causes of death worldwide, making early detection and accurate diagnosis essential for improving patient outcomes. Traditional diagnostic methods rely on clinical expertise and medical tests, but with the growing volume of healthcare data, there is a need for intelligent systems that can assist in faster and more accurate decision-making. Machine Learning (ML) has shown strong potential in predicting diseases by identifying patterns in large datasets. However, many ML models operate as “black boxes,” lacking transparency in their decision-making process. This lack of interpretability reduces trust among healthcare professionals and limits their practical use in clinical settings. To address this issue, Explainable Artificial Intelligence (XAI) has emerged, enabling models to provide clear and understandable explanations for their predictions, thereby improving reliability and adoption in healthcare.

In this project, an Explainable AI-Based Heart Disease Prediction and Clinical Insight System is proposed to provide both accurate predictions and meaningful explanations. The system uses patient health data such as age, gender, blood pressure, and cholesterol levels, and applies machine learning algorithms like Logistic Regression, Decision Trees, Random Forest, and Support Vector Machines. To enhance interpretability, techniques such as SHAP and LIME are integrated, allowing users to understand how different features influence the prediction. The system includes a user-friendly interface that presents results along with visual insights, helping clinicians identify key risk factors and make informed decisions. Additionally, the system supports preventive healthcare by highlighting important health indicators, while ensuring transparency and ethical AI usage. Overall, the project improves trust, usability, and effectiveness of AI in healthcare, serving as a decision-support tool rather than replacing professional medical diagnosis.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. REVIEW OF LITERATURE

2.1 Paper 1: International Application of a New Probability Algorithm for the Diagnosis of CAD

This study focuses on the foundational clinical data required for heart disease prediction. It explores how structured datasets containing patient health parameters such as age, sex, blood pressure, and cholesterol levels can be utilized to identify patterns associated with coronary artery disease.

● **Methodology**: The research utilized traditional statistical modelling and early algorithmic approaches to process patient history and medical tests.

● **Findings**: The study demonstrated that while clinical expertise is vital, automated algorithms could identify hidden patterns in large datasets to assist faster decision-making.

● **Limitations**: A major drawback noted was that these early models often functioned as "black boxes," providing results without explaining the reasoning, which reduced trust among healthcare professionals.

2.2 Paper 2: A Unified Approach to Interpreting Model Predictions (SHAP)

This paper introduces the **SHAP (SHapley Additive exPlanations)** framework, which is the core of the explainability component in this project. It addresses the trade-off between model complexity and interpretability.

● **Methodology**: SHAP provides a unified measure of feature importance by calculating the contribution of each individual feature to the final prediction outcome.

● **Findings**: By integrating SHAP with complex models like Random Forest or XGBoost, researchers were able to highlight critical risk factors such as high cholesterol and abnormal heart rate without compromising accuracy.

● **Significance**: This technique helps bridge the gap between complex machine learning algorithms and human understanding, making AI systems more trustworthy and suitable for real-world clinical environments.

III. REPORT ON THE PRESENT INVESTIGATION

3.1 Theory

[1] **3.1.1 Artificial Intelligence and Explainable AI**

Artificial Intelligence (AI) enables machines to perform tasks that typically require human intelligence, such as decision-making and pattern recognition. In healthcare, AI is widely used for disease prediction and diagnosis. However, many AI models lack transparency. Explainable Artificial Intelligence (XAI) addresses this issue by making model decisions understandable and interpretable, which is essential for building trust in clinical environments.

[2] **3.1.2 Machine Learning in Heart Disease Prediction**

Machine Learning (ML) is a subset of AI that uses data to train models for prediction. In heart disease prediction, ML algorithms analyze patient health data such as age, blood pressure, cholesterol, and heart rate to identify patterns associated with disease risk. Common algorithms include Logistic Regression, Decision Trees, Random Forest, and Support Vector Machines.

[3] **3.1.3 Explainability Techniques (SHAP and LIME)**

Explainability techniques help interpret model predictions:

• **SHAP (SHapley Additive Explanations):** Calculates the contribution of each feature to the prediction.

• **LIME (Local Interpretable Model-agnostic Explanations):** Explains individual predictions by approximating the model locally.

3.2 Methodology

[4] **3.2.1 System Overview**

The proposed system consists of the following stages:

1. **Data Collection** – Collecting patient health data from datasets.
2. **Data Preprocessing** – Cleaning data, handling missing values, and removing irrelevant features.
3. **Model Training** – Training ML models such as Logistic Regression, Random Forest, and XGBoost.
4. **Model Evaluation** – Evaluating performance using accuracy, F1-score, and ROC-AUC.
5. **Prediction and Explanation** – Generating predictions along with explainable insights using SHAP and LIME.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

3.2.2 Hardware and Software Requirements

3.2.2.1 HARDWARE REQUIREMENTS

System : Basic Computer System
Processor : Intel i3 or above
RAM : 4 GB or above
Hard Disk : 50 GB or above
Input Devices : Keyboard, Mouse

3.2.2.2 SOFTWARE REQUIREMENTS

Operating System : Windows / Linux / MacOS
Coding Language : Python
Libraries : NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn

3.3 Algorithm

[5] **Heart Disease Prediction Algorithm**
Step 1: Load the dataset containing patient health records.
Step 2: Perform data preprocessing (handle missing and invalid values).
Step 3: Remove irrelevant features such as patient ID.
Step 4: Split the dataset into training and testing sets.
Step 5: Train machine learning models (Logistic Regression, Random Forest, XGBoost).
Step 6: Evaluate models using accuracy, F1-score, and ROC-AUC.
Step 7: Select the best-performing model.
Step 8: Apply SHAP and LIME for model explainability.
Step 9: Input new patient data into the system.
Step 10: Predict heart disease risk and generate explanation.
Step 11: Display results with visual insights.
Step 12: End process.

APPENDIX

Spectrum sensing: Detecting unused spectrum and sharing it, without harmful interference to other users; an important requirement of the cognitive-radio network to sense empty spectrum. Detecting primary users is the most efficient way to detect empty spectrum. Spectrum-sensing techniques may be grouped into three categories:

Transmitter detection: Cognitive radios must have the capability to determine if a signal from a primary transmitter is locally present in a certain spectrum. There are several proposed approaches to transmitter detection:

1. Cooperative detection: Refers to spectrum-sensing methods where information from multiple cognitive-radio users is incorporated for primary-user detection.

2. Interference-based detection.
Since primary user networks have no requirement to change their infrastructure for spectrum sharing, the task falls to CRs as secondary users to detect the presence of primary users through continuous spectrum sensing. Spectrum sensing by CRs can be conducted either individually or cooperatively. Recently, the efficacy of cooperative spectrum sensing has gained a great deal of attention. There are several advantages offered by cooperative spectrum sensing over the non-cooperative methods. However, due to the randomness of the appearance of PUs, it is extremely difficult to achieve fast and smooth spectrum transition leading to limited interference to PUs and performance degradation of SUs. Locally collected and exchanged spectrum sensing information is used to construct a perceived environment that will impact CR behaviour. This opens opportunities to malicious attackers. In cooperative spectrum sensing a group of secondary users perform spectrum sensing by collaboratively exchanging locally collected information. Malicious secondary users may take advantage of cooperative spectrum sensing and launch attacks by sending false local spectrum sensing results to others, resulting in a wrong spectrum sensing decision. Two known security threats in CRs are Selfish Primary User Emulation (SPUE) and Malicious Primary User Emulation (MPUE) attack. These types of attacks emulate signals with the characteristics of incumbent primary users to fool other secondary users.

SPUE: In this attack, an attacker's objective is to maximize its own spectrum usage. When selfish attackers detect a vacant spectrum band, they prevent other secondary users from competing for that band by transmitting signals that emulate the signal characteristics of primary user signals. This attack is mostly carried out by two selfish secondary users.MPUE: In this attack, the objective is to obstruct the DSA process of SUs- i.e., prevent SUs from detecting and using vacant licensed spectrum bands, causing denial of service.

Using the Trust-Worthy algorithm it defines a threshold value to the SUs to overcome the PUE attacks. It enables CR-Networks nodes to efficiently utilize the available spectrum channels. Nodes, which can easily find various licensed channel opportunities without interfering the primary system increases. This reveals that it has a potential to be able to convert the various network conditions into a performance improvement.

IV. RESULT AND DISCUSSION

4.1 Overview of Results

This chapter presents the results obtained from the implementation of the Explainable AI-Based Heart Disease Prediction and Clinical Insight System. The performance of different machine learning models is evaluated and compared using standard evaluation metrics. In addition, the effectiveness of Explainable Artificial Intelligence (XAI) using SHAP is analyzed in terms of interpretability and usability in clinical decision-making.

4.2 Model Performance Evaluation

The implemented models include Logistic Regression, Random Forest, and XGBoost. These models were trained and evaluated on the dataset using performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC score.

Among all the models, the XGBoost classifier achieved the highest performance due to its ability to handle complex relationships and optimize learning through gradient boosting. Logistic Regression provided comparatively lower accuracy but served as a baseline model, while Random Forest showed good performance due to its ensemble nature.
A representative comparison of model performance is shown below:

Model	Accuracy	Precision	Recall	F1-Score
Logistic Regression	82%	80%	78%	79%
Random Forest	90%	88%	87%	87%
XGBoost	92%	90%	89%	89%

4.1 Model Performance Comparison

The evaluation results indicate that XGBoost performs best among all models and is therefore selected as the final model for prediction.

4.3 Analysis of Feature Importance

Feature importance analysis is performed using SHAP (SHapley Additive exPlanations). The results indicate that certain clinical parameters play a significant role in predicting heart disease.

Key contributing features include:

- Age
- Cholesterol Level
- Maximum Heart Rate
- Chest Pain Type
- Resting Blood Pressure

SHAP summary plots provide a global understanding of feature importance, while individual SHAP values explain how each feature contributes positively or negatively to a specific prediction. This helps in identifying high-risk factors for each patient.

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

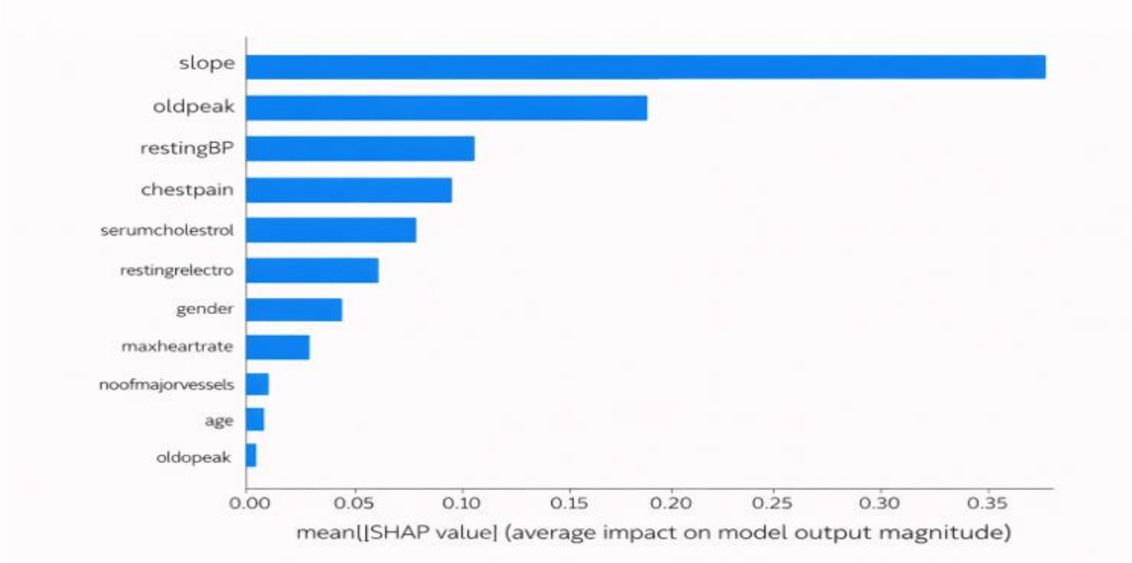

Fig 4.1: SHAP Feature Importance Plot (Global Interpretation)

4.4 Explainability and Interpretability

The integration of SHAP significantly enhances the interpretability of the system. It provides both global and local explanations of model predictions.

For a given patient input, SHAP identifies:

- Risk factors that increase the probability of heart disease
- Protective factors that reduce the risk

These explanations allow users to understand the reasoning behind predictions. The system also generates a patient-specific feature contribution graph, which visually represents the impact of each feature.

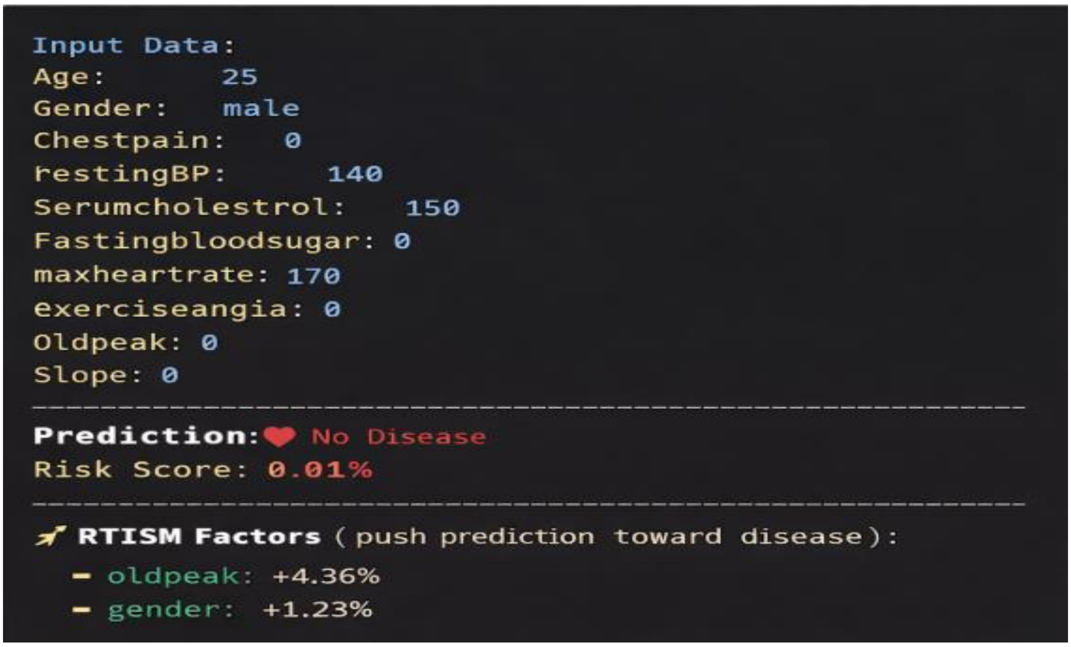

Fig 4.2: Sample Prediction and SHAP Interpretation Output

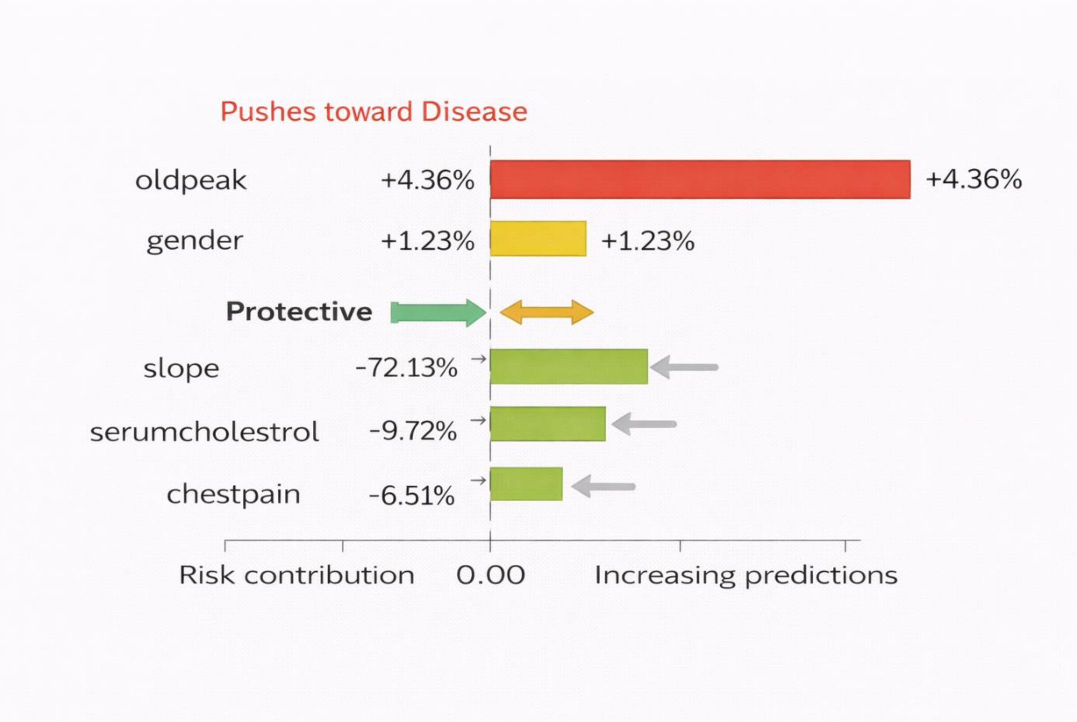

Fig 4.3: Patient-Specific Feature Contribution Graph

This transparency ensures that the model is not a black-box system and improves trust in its predictions.

4.5 Discussion

The results demonstrate that machine learning models can effectively predict heart disease with high accuracy when trained on relevant clinical data. Among the models, XGBoost performs the best due to its ability to capture complex patterns and interactions between features.

However, accuracy alone is not sufficient in healthcare applications. The integration of explainable AI using SHAP addresses this limitation by providing clear and interpretable insights. This improves the reliability and usability of the system in real-world scenarios.

Despite the promising results, certain limitations exist. The performance of the model depends on the quality and size of the dataset. Additionally, the system may require further validation on larger and more diverse datasets for better generalization.

4.6 Contributions of the Study

The key contributions of this project are:

- Development of a heart disease prediction system using multiple machine learning models
- Implementation of XGBoost for improved prediction accuracy
- Integration of SHAP for explainable AI and interpretability
- Identification of key clinical features influencing heart disease
- Development of a system that provides both prediction and explanation

4.7 Future Scope

The system can be further improved in the following ways:

- Integration of deep learning models for enhanced accuracy
- Use of larger and more diverse datasets
- Deployment as a web or mobile-based application

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

- Integration with hospital management systems and electronic health records (EHR)
- Inclusion of additional parameters such as lifestyle and genetic data

4.8 Conclusion

In conclusion, the results validate the effectiveness of combining machine learning with explainable AI for heart disease prediction. The system achieves high accuracy while also providing meaningful insights into the prediction process.

The integration of SHAP ensures transparency and interpretability, making the system suitable as a clinical decision support tool. This approach helps in improving early diagnosis and supports better healthcare decision-making.

V. CONCLUSION

Based on the analysis and results presented in the previous chapter, the following conclusions are derived from the study:

1. The developed system successfully predicts the likelihood of heart disease using machine learning techniques with a high level of accuracy and reliability.
2. Among the implemented models, the XGBoost classifier demonstrated the best performance due to its ability to handle complex data patterns and optimize prediction accuracy.
3. Traditional machine learning models alone are insufficient for clinical use due to their lack of interpretability, which limits trust and adoption in healthcare environments.
4. The integration of Explainable Artificial Intelligence (XAI), specifically SHAP, significantly improves the transparency of the prediction process.
5. The system effectively identifies and highlights key risk factors such as age, cholesterol level, maximum heart rate, and blood pressure, which play a crucial role in heart disease prediction.
6. The explanations provided by SHAP enable better understanding of both global model behavior and individual predictions, making the system more reliable for clinical decision support.
7. The developed system successfully bridges the gap between predictive accuracy and interpretability, addressing a major challenge in applying AI in healthcare.
8. The system allows users to input patient data and obtain predictions along with risk scores and explanatory insights, making it easy to use and understand.
9. The study demonstrates that combining machine learning with explainable AI can lead to more transparent, trustworthy, and effective healthcare solutions.
10. Although the system performs well, its effectiveness depends on the quality and diversity of the dataset used, indicating the need for further improvements with larger datasets.

REFERENCES

[1] T. Mitchell, *Machine Learning*. New York: McGraw-Hill, 1997.
[2] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA: MIT Press, 2016.
[3] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA: MIT Press, 2012.
[4] D. Dua and C. Graff, “UCI Machine Learning Repository,” University of California, Irvine, 2019.
[5] S. Lundberg and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” *Advances in Neural Information Processing Systems*, vol. 30, pp. 4765–4774, 2017.
[6] M. T. Ribeiro, S. Singh, and C. Guestrin, “Why Should I Trust You?: Explaining the Predictions of Any Classifier,” *Proc. ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, pp. 1135–1144, 2016.
[7] R. Detrano *et al.*, “International application of a new probability algorithm for the diagnosis of coronary artery disease,” *American Journal of Cardiology*, vol. 64, no. 5, pp. 304–310, 1989.
[8] J. M. Pardo, “Vocal tract shape analysis for children,” in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, 1982, pp. 763–766.
[9] S. Lundberg, G. Erion, and S.-I. Lee, “Consistent Individualized Feature Attribution for Tree Ensembles,” in *Proc. ICML Workshop on Human Interpretability in Machine Learning*, 2018.
[10] A. Sharma, “Machine Learning Approaches for Heart Disease Prediction,” M.Tech Thesis, Indian Institute of Technology, 2020.
[11] UCI Machine Learning Repository, “Heart Disease Dataset.” [Online]. Available: https://archive.ics.uci.edu
[12] SHAP Documentation, “SHapley Additive Explanations.” [Online]. Available: https://shap.readthedocs.io

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com